Impact of Spatial Competition and Agglomeration Effects on the Temporal Transferability of Destination Choice Models

A Case Study on the Metropolitan Region of Amsterdam

by

Rohan Menezes

in partial fulfillment of the requirements for the degree of

Master of Science

in Transport, Infrastructure and Logistics
at the Delft University of Technology,

to be defended publicly on Wednesday, December 18, 2024 at 10:30 AM.

Student number:	5850908	
Thesis committee:	Dr. Ir. G. (Gonçalo) Correia	TU-Delft, Committee Chair
	Dr. Ir. M. (Maarten) Kroesen	TU-Delft, Supervisor
	Dr. Ir. H. (Han) Zhou	TNO, External Supervisor

Preface

This thesis marks the end of my incredible experience of a Master's in Transport Infrastructure and Logistics at TU Delft. In this thesis, I have touched upon the theme of the essence of modeling and why, rather than performance, it is crucial to focus on creating models that accurately represent the systems they are intended to be simplified replicas of. While here I touch upon choice modeling, and therefore the focus is on enabling models to better represent behavior, it applies to modeling in general. I discovered that sometimes the answers are simple—you just need to focus on the essence and look back every now and then for solutions.

At TU Delft, I discovered my competency and love for modeling systems. I used to hate mathematics. Avoid coding like a plague. and now I am fascinated by them. Thanks, Yousef, for making me realize how beautiful mathematics can be. I have dabbled in mathematical modeling, an MPC (Model Predictive Control) system, and attended simulation classes early in the mornings just out of curiosity. I learned that my strength lies in versatility and not in specializing in any one way of modeling. By sheer luck, TIL was a perfect match for who I was. providing me a platform to indulge in all of it. I am grateful I met some incredibly intelligent people, be it teachers or my classmates, who were patient enough to help me learn. I am grateful for their kindness.

This Master's and this thesis were full of challenges. During this two-and-a-half-year journey, I have failed countless times. But all these challenges, which have often defeated me, have made me more grateful and reconditioned me to have a new, more resilient approach. It reminds me of the poem by Rainer Maria Rilke, "The Man Watching." The thing about failing so many times is that eventually, you get used to getting up. You learn to dust yourself off and begin again. Yes, failure hurts, but I am starting to no longer dread it. I see it as an opportunity to do the one thing I always enjoyed: Learn.

Gonçalo, thank you for being critical and supportive, for being patient enough to teach me through my mistakes. Maarten, thank you for being kind. Your words of encouragement always helped me get back to believing in myself whenever I felt doubtful of my abilities. Han, thank you for supporting me all the times when I felt stuck. Thank you, Jip Classens; without your constant support and patience in answering all my questions about the travel time calculations, I would never have completed this thesis.

Without my closest people, facing these challenges alone would have been very isolating. Thank you, Fariba, John, Inez, Tejas, Aparna, Avishya, and Axin, for being there. I have many more people I met here to thank, but unfortunately, the page length does not allow me. To all the people back home, my school friends; thank you for keeping on believing in me. Tej, thank you for being so supportive of me. Mom, I know there were times when everything felt confusing, especially with us communicating over such a long distance. Without you, I would never have been able to enjoy this privileged experience of learning. Thank you for allowing me to express myself, especially when it was difficult to comprehend the situation.

I am grateful to TU Delft. I learned how to unlearn and learn. and how to apply myself. I have learned more than I could ever imagine. I have become a problem solver. From a workshop at TU Delft, I love this quote: "You are a human being, not a human doing." More than solving problems, more than being a student, I have realized that learning makes me want to live, experience this world and its beauty, and take it all in. No one aspect defines you. I now realize this thesis, too, touches upon a similar theme. Funny.

In life, I will face many challenges, but this experience has given me the tools to navigate through them. I am grateful I will soon be a TU Delft alumnus. Perhaps that's the hallmark of a great education system—it teaches you things that prepare you to live a good life. I am glad I grabbed this opportunity with both hands. Thank you, TU Delft.

Rohan Menezes
Den Haag, December 2024

Summary

Transport planning relies heavily on models for forecasting travel behavior, owing to the long-term impact of policies and their resource-extensive execution. To ensure transport models are effective tools for planning, they must not only adequately explain current travel choices but also maintain predictive accuracy over forecast horizons while being computationally feasible for practical use. A simple approach to achieving this is by including behavioral theories in transport models to improve the behavioral representation of models. Although these theories may enhance the model's ability to explain current travel choices, they do not necessarily improve their forecasting abilities due to the risk of overfitting. Overfitted models tend to explain random noise rather than the signal in the data (Parady et al., 2021). Thus, overfitting negatively impacts the model's temporal transferability, i.e., its ability to maintain predictive accuracy across forecast horizons.

One such theory is Spatial Competition and Agglomeration Effects (SC&AE). This theory examines how opportunities present nearby influence the attractiveness of a destination. This influence can be either positive (Agglomeration) or negative (Spatial Competition). Although SC&AE is widely recognized in the literature for enhancing the explanatory power of computationally simple Multinomial Logit (MNL) destination choice models for various trips, its impact on the temporal transferability of these models remains unexplored.
This study assesses the impact of SC&AE on temporal transferability of MNL destination choice models for home-based maintenance, work, and education trips in the Metropolitan Region of Amsterdam on a 5-year short-term forecast horizon (2018-2022). By providing quantitative evidence of the impact on temporal transferability, this research aims to assess the validity of SC&AE in explaining destination choices and thus justify its inclusion in MNL DCMs.

In order to assess this impact, four datasets for 2018 and 2022, such as the Dutch National travel survey (ODiN), employment, enrollment, and travel time matrix for the MRA region, are used. Based on the findings of the literature review and information available in the data, variables are selected, and data is processed to facilitate the estimation of parameters for 2018 and 2022 across the three trip purposes. The impact is then assessed by comparing the performance of an MNL DCM specification that excludes SC&AE effects with the model specification that includes them. For the assessment, four performance indicators across three categories of statistical(Transfer Index), predictive (%Correct Predictions) , and quality of probabilistic predictions (Fitting Factor, Prediction Clearness) in order to gain insights into the key aspects of the destination choice model in terms of impact. Since sampling often accompanies destination choice models, the performance on these indicators is also compared across full, random, and a variant of stratified importance sampling to understand the impact of SC&AE on temporal transferability varies with sampling methods.

Consistent with previous research, it finds statistically significant negative SC&AE parameters for home-based maintenance (HBM) and work trips in 2018 and 2022, indicating a dominant spatial competition. HBM trips exhibited more negative SC&AE parameters (-1.60 to -1.77) than work trips (-0.67 to -0.76), reflecting a stronger influence on destination choices due to easily substitutable destinations and high ease of switching. In contrast, work trips involve destinations with fewer alternatives nearby and longer-term commitments, resulting in a lesser influence of spatial competition.
For secondary and above education level trips, the positive SC&AE parameter suggests that agglomeration effects dominate spatial competition. Unlike Sá et al. (2004), who found dominating spatial competition in university choices among high school graduates in the Netherlands by focusing solely on one post-secondary education level, this study included both secondary and multiple higher education levels. Hence, the broader range of nearby educational options to continue education makes areas with a higher number of institutions more attractive, resulting in a positive value that reflects dominant agglomeration effects. For primary education trips, the SC&AE parameter in 2018 was statistically insignificant, indicating a negligible influence on destination choices. This could be due to young children having the highest commitment period and the lowest flexibility to switch schools. Hence, the spatial distribution of primary school opportunities has little effect on their location choices.

Additionally, this study extends beyond evaluating the explanatory power of SC&AE in a single context by exploring its contribution to the temporal transferability of destination choice models. The findings indicate

that SC&AE has a positive but limited impact on temporal transferability, varying by trip purpose: highest for home-based maintenance (HBM) trips, followed by work trips, and inconsistent for secondary & higher education (positive on the Transfer Index but negative on rest of the indicators).

A closer examination of log-likelihoods and other performance indicators (such as fitting factor, percentage of correct predictions, and clearness of predictions) confirms that the actual impact is limited. This exaggeration occurs because TI relies solely on the ratio of gains in log-likelihoods (LL), which can misrepresent models with small absolute gains. For example, a model with an LL gain ratio $\frac{1}{2}$ and another with a gain ratio of $\frac{50}{100}$ will have the same TI value of 0.5. But clearly, the second model is much better and will perform positively on other indicators, while the model with LL gain ratio $\frac{1}{2}$ will perform poorly on other indicators
Therefore, TI values should always be presented alongside other performance measures or at least be accompanied by a comparative analysis of the log-likelihood values used in their calculation.

Observing the trend of varying impacts of SC&AE on temporal transferability by trip purpose, the impact of SC&AE on temporal transferability decreases with decreasing traveler autonomy and ease of switching destinations: it is highest for HBM trips, where travelers have high autonomy and flexibility, less so for work trips due to longer commitment periods, and inconsistent (positive on the Transfer Index but negative on rest of the indicators) for secondary and above level education trips, where travelers are effectively committed to institutions until they complete their education. Considering the low amount of trips for secondary and above education trips compared to other trips, these inconsistent results for secondary and above education trips should be considered inconclusive. However, considering autonomy and ease of switching, the impact is likely to be limited, more so than for the HBM and work trips. For primary education trips, autonomy and flexibility are the lowest, thus explaining the statistically insignificant estimated SC&AE parameter for primary education trips.
Moreover, the choice of sampling method affects the temporal transferability of destination choice models and, thus, the impact of SC&AE on temporal transferability too. Models using Stratified Importance Sampling (SIS) show higher TI values than those using Random Sampling (RS), with and without the SC&AE parameter. While the initial performance boost from SIS reduces the absolute gain in the TI value from including SC&AE compared to RS, SIS allows models to achieve higher overall TI values after including SC&AE.
Overall, Given the minimal effort required to include these effects in an MNL model because it reuses existing information such as zonal size measures and travel impedance, SC&AE provides technically "free" robust log-likelihood gains.

From a policy perspective, including SC&AE into destination choice models enhances their effectiveness as predictive tools by improving their temporal transferability. Models that include SC&AE maintain predictive accuracy over time, which is particularly valuable for scenarios where travelers have significant autonomy and flexibility, such as discretionary activities such as shopping and maintenance trips. This added robustness stems from SC&AE's ability to address two fundamental flaws of Multinomial Logit (MNL) models that limit their behavioral accuracy in representing travelers' destination choices: (1) the Independence of Irrelevant Alternatives (IIA) assumption, and (2) neglecting the influence of the spatial distribution of opportunities. By including information about all alternative destinations, SC&AE tackles both issues, allowing MNL models to become more behaviorally representative whilst retaining their computational simplicity.

The results of this study have broader implications for transport modeling. To overcome the limitations of MNL models, researchers have often relied on more complex disaggregate models, which are computationally intensive and often impractical for large datasets. However, this study demonstrates that simpler models such as MNL can overcome their flaws by integrating theories such as SC&AE, thereby improving behavioral representation while retaining computational efficiency. By focusing solely on whether a model explains or predicts behavior well, we may have been asking the wrong questions. Instead, we should ask, "Is my model an accurate representation of the system it is supposed to represent?" By addressing this fundamental question and bringing models closer to accurately reflecting the system, we automatically enhance their explanatory and predictive capabilities.
This research shows that we do not necessarily need to rely on more complex models; there is another simpler way: using theories to enhance simpler models. When these simple models reach their limit on how much they can be improved using behavioral theories, which they eventually will, this approach might allow another pathway to improve more complex models, making them more computationally feasible and data-efficient without relying primarily on advancements in computational science. To achieve this, we need to look beyond the transport domain and draw insights from related fields, such as psychology or other behavioral sciences. It may be time for transport modelers to look beyond their transport domain and integrate psychological theories to make transport models a better representation of how travelers make choices. The results of this study are certainly encouraging.

Abbreviations

Abbreviation	Meaning
ABM	Activity-Based Modelling
BART	Bay Area Rapid Transit
BRON	Basis Register Onderwijs (Dutch Official Educational Database)
BRP	Basisregistratie Personen (Dutch Personal Records Database)
CBS	Centraal Bureau voor de Statistiek (Dutch Central Bureau of Statistics)
DCM	Destination Choice Model
FF	Fitting Factor
GDPR	General Data Protection Regulation
GM	Gravity Model (Disaggregate equivalent GM for Transfer Index)
HAVO	Hoger Algemeen Voortgezet Onderwijs (Higher General Secondary Education)
HBM	Home-Based Maintenance
HBO	Hoger Beroepsonderwijs (Higher Professional Education)
IIA	Independence of Irrelevant Alternatives
IVTT	In-Vehicle Travel Time
LoS	Level of Service
MBO	Middelbaar Beroepsonderwijs (Middle-Level Vocational Education)
MNL	Multinomial Logit
MRA	Metropolitan Region of Amsterdam
ODiN	Onderzoek Verplaatsingen in Nederland (Dutch National Travel Survey)
OLS	Ordinary Least Squares
OVTT	Out-of-Vehicle Travel Time
OViN	Onderzoek Verplaatsingen in Nederland (Former Dutch National Travel Survey)
PC4	Four-digit postal code
SCAE	Spatial Competition and Agglomeration Effects
SES	Socioeconomic Status
SQL	Structured Query Language
TI	Transfer Index
TNO	Dutch Organization for Applied Scientific Research
VMBO	Voorbereidend Middelbaar Beroepsonderwijs (Preparatory Secondary Vocational Education)
VWO	Voorbereidend Wetenschappelijk Onderwijs (Pre-University Education)
WO	Wetenschappelijk Onderwijs (University Education)

Contents

List of Figures

List of Tables

1
Introduction

1.1. Problem Statement and Research Aim

Transport planning relies heavily on models for forecasting travel behavior, owing to the long-term impact of policies and their resource-extensive execution. Government agencies at various levels use strategic forecasting models to predict demand for current and planned transport infrastructure and to evaluate the effectiveness of various policy options (Fox et al., 2014). Hence, while it is important to improve models that can better explain current travel choices, it is far more crucial that they also increase their ability to maintain predictive accuracy across forecast horizons.

One strategy for improving transport models involves developing theories that aid us in understanding travel behavior better. Although these theories may enhance the model's ability to explain current travel choices, they do not necessarily improve their forecasting abilities. This is because of the risk of overfitting, where the model is fitted to random noise rather than the signal in the data (Parady et al., 2021). This negatively impacts the model's temporal transferability i.e., its ability to maintain predictive accuracy across forecast horizons.

One such theory is Spatial Competition and Agglomeration Effects (SC&AE). This theory examines how opportunities present nearby influence the attractiveness of a destination. This influence can be either positive (Agglomeration) or negative (Spatial Competition). By accounting for this spatial heterogeneity, it remedies the popular Independence of Irrelevant Alternatives (IIA) assumption of Multinomial Logit (MNL) destination choice models (DCM). This substantially improves the model's explanatory power for current destination choices. Specifically, this improvement occurs when SC&AE are accounted for through accessibility measures in the utility specification. This is well noted in transport literature across various trip purposes and geographical regions over the years (Bernardin et al. (2009); Ho and Hensher (2016); Sá et al. (2004)).

Thus, given its long-term applicability in various contexts, these effects are potentially essential in explaining traveler's destination choices and thus can contribute significantly to DCM's temporal transferability. However, the impact of SC&AE on the temporal transferability of DCMs has remained unexplored.

Using a case study approach in the Metropolitan Region of Amsterdam (MRA), focusing on work, education, and maintenance trips, this study compares the performance of an MNL DCM specification that excludes SC&AE effects with the model specification that includes them. Taking into consideration ODiN travel survey data availability, the comparison is made across statistical and predictive measures during a 5-year short-term forecasting horizon (2018-2022). This approach sheds light on how SC&AE influences travelers' destination choices and how it varies by trip purpose within the MRA. This research is conducted in collaboration with TNO, the Dutch Organization for Applied Scientific Research.

To sum up, this research assesses the validity of SC&AE in explaining destination choices to justify its inclusion in MNL DCMs. It does so, by providing quantitative evidence of the SC&AE impact on the temporal transferability of DCMs. This enables assessing whether SC&AE captures travel behavior that drives destination choices rather than merely capturing behavior contextually in the travel survey data. Additionally, this evidence enables TNO researchers to compare the effort and complexity of integrating SC&AE effects into the model development process against the improvements in the model's temporal transferability. Moreover, the methodology could be extended into an experimental framework for comprehensive scenario analysis. This would help draw comparisons between different model specifications, enabling TNO to develop more robust and effective DCMs in the future.

1.2. Research Scope

Since TNO intends to implement these DCMs in the Activity-Based Modelling (ABM) framework[1] in the future, This research focuses on three trip purposes to align with the ABM framework and TNO's objectives:

- Work
- Education
- Home-Based Maintenance Trips (Shopping, Personal care)

The study area for this research is the Metropolitan Region of Amsterdam (MRA). It consists of 30 municipalities (See Appendix D). As of 2023, approximately 2.5 million individuals, accounting for over 14% of the Netherlands' population reside in the MRA. This region is one of the country's strongest economic areas and holds substantial international importance (Metropoolregio Amsterdam, 2023). As a result, this study area is crucial for TNO, and they possess extensive traffic data for further validation and conduct ongoing transport research in this area. Additionally, they intend to use the models developed during this research, to test the model's spatial transferability, i.e., how well the model maintains its predictive ability when applied to different geographical regions (Parady et al., 2021). Given the MRA's economic and population diversity, it is an ideal study area for developing DCMs with potentially high spatial transferability. These models can provide valuable insights applicable to MRA and significant implications for transport planning and policy-making in urban areas throughout the Netherlands and globally.

Furthermore, the forecasting period chosen for this study is 5-year short-term forecasting horizon (2018-2022). This forecasting horizon period was chosen due to data availability constraints concerning the Dutch National Travel survey data (ODiN), one of the data sources used for developing and estimating the destination choice models for the three trip purposes.
ODiN is a continuation of the "Onderzoek Verplaatsingen in Nederland" (OViN) survey, which was carried out by CBS from 2010 to 2017. The methodology of ODiN 2018 differs significantly from that of the earlier OViN survey, creating a methodological break between the two. As a result, the findings from ODiN cannot be directly compared with those from OViN analyses (DANS, 2024). ODiN surveys only individuals aged six and older, excluding younger children, which reduces the target population by over 1 million. It uses an internet-only (cawi) data collection method, unlike OViN's mixed-mode approach. ODiN also integrates all domestic and international vacation trips as regular trips, whereas OViN handles these trips separately. In addition, ODiN determines the primary mode of transport based on distance, unlike OViN's priority-based system. Finally, ODiN utilizes more register-based data called Basisregistratie Personen (BRP), reducing the need for extensive survey questions (Statistics Netherlands (CBS), 2019).

Although adjustments could improve the comparability between ODiN and OViN, some differences would likely remain due to inherent methodological changes. Given the limited timeframe, making such modifications is beyond the scope of this study. Therefore, considering that the earliest available ODiN data is from 2018 and the most recent from 2022, a short-term forecasting horizon of five years (2018-2022) is selected for this analysis.

[1] See Appendix A for the Activity Based Modelling framework.

1.3. Research Questions

By providing quantitative evidence of the impact on temporal transferability, this research aims to assess the validity of SC&AE in explaining destination choices and thus justify its inclusion in MNL DCMs. The main research question that addresses this aim is as follows:

How do SC&AE affect the temporal transferability of destination choice models ?

This main question will be answered by addressing the following sequence of sub-questions:

1. What factors affect the temporal transferability of destination choice models?
2. How do existing destination choice models incorporate SC&AE and what is the significance of these effects on destination choices?
3. How do SC&AE impact the destination choices of travelers in the study area and how do they differ across trip purposes?
4. How do the DCMs for various trip purposes perform across statistical tests and predictive measures and what are their policy implications for the study area?

1.4. Research Approach

Table 1.1 summarizes the methodologies used to answer the sub-research questions and the sections of the document where they are addressed. Overall, there are three methods used to answer the sub-questions: Literature Review, Discrete Choice Modelling, and Validation.

Table 1.1: Research Method per research question

Sub-RQ	Method	Detail	Section
1	Literature Review	Identify factors influencing the temporal transferability of transportation models. Use the factors to motivate the selection of variables to be included in the to-be-developed destination choice models.	Section 2.4
2	Literature Review	Evaluate existing destination choice models that consider SCAE and other explanatory variables to motivate the structure of the SCAE and other explanatory variables when developing the model specification.	Section 2.3
3	Discrete choice modelling, Results Analysis	Based on the findings of the above two sub-questions and information in the data, use the selected variables to develop model specifications for the three trip purposes. Estimate the parameters of the MNL models, examine them, and compare the parameter estimates across different trip purposes.	Section 4.6, Chapter 5
4	Validation tests, Results Analysis	To run statistical, predictive, and quality probabilistic prediction performance indicators and draw insights from these results. Use the insights to reflect on SCAE's role in increasing the effectiveness of destination choice models and discuss the broader implications for developing transport models as a reliable tool for transport policy planning.	Chapter 5, Section 5.6

1.5. Thesis Outline

This thesis is structured as follows: Chapter 2 reviews the literature to introduce the concepts of Multinomial Logit destination choice models, temporal transferability, Spatial Competition and Agglomeration effects, and then answers the first two research questions. Chapter 3 provides details regarding the data sources used to estimate the models. Chapter 4 details the methodology used in this study. Chapter 5 provides the analysis and discusses the implications of the results. Chapter 6 summarizes the answers to the research questions, details the limitations of this thesis, and provides recommendations and possible directions for future research.

2

Literature Review

The literature review is organized into several key sections, each providing a comprehensive overview of the theories, models, and methodologies relevant to this research.

Section 2.1 focuses on Multinomial Logit (MNL) models for destination choice. It outlines the basic structure of these models. The section addresses key limitations of traditional MNL models, such as the Independence of Irrelevant Alternatives (IIA) assumption, and reviews how previous research has integrated SC&AE to overcome this IIA assumption to make destination choice predictions more realistic.

Section 2.2 introduces the concept of Spatial competition and Agglomeration effects (SC&AE). It highlights how the proximity of opportunities can influence the attractiveness of destinations, using examples to demonstrate these effects.

Section 2.3 examines studies that include SC&AE in destination choice models for the three trip purposes relevant to this research. Additionally, it reviews studies on educational location choices in the Netherlands that do not include SC&AE, to understand how other variables are used to explain these choices within the Dutch education system.

Section 2.4 elaborates on the concept of temporal transferability, explaining its importance in transport models and how it is assessed through external validation. Previous studies on the temporal transferability of mode and mode-destination choice models are reviewed to identify key factors that affect a choice model's temporal transferability.

Lastly, Section 2.5 concludes the literature review by identifying research gaps and the aim of this research. It does so by discussing the risks of overfitting and its negative impact on a model's predictive power. It emphasizes the importance of balancing explanatory and predictive accuracy, the significance of including SC&AE into DCMs, and how this integration can improve the models' temporal transferability.

2.1. Multinomial Logit models of Destination Choice

Trip Distribution Models are a fundamental sub-component of a larger transportation modeling framework. They are used to predict how trips generated at one location (origin) are distributed to other locations (destinations) within a region (de Dios Ortúzar and Willumsen, 2011). Essentially, these models estimate the flow of trips between different zones as a function of travel impedance factors such as distance, travel cost, etc., which deter the flow of trips, and trip type-dependent destination attractiveness factors such as employment, population, etc. that have a positive influence on the flow of trips to the destinations.
Trip distribution models can be broadly classified as aggregate and disaggregate models. Aggregate models distribute trips based on observed patterns for groups of travelers or average relationships at the zonal level. One common aggregate method is the gravity model, which draws an analogy to Newton's law of gravitation. In this model, trip flows are proportional to the product of origin and destination attractiveness and inversely proportional to travel impedance (de Dios Ortúzar and Willumsen, 2011). The gravity model, which is mathematically simple and computationally efficient, became the most commonly used model for transport planning (Bernardin et al., 2009). Disaggregate models, on the other hand, are based on observed choices at the individual traveler or

household level. The most common type of disaggregate model is the discrete choice model based on Utility Maximization theory. Because these disaggregate models are based on theories of individual behavior and do not rely on physical analogies, they have a potential advantage in that they are more likely to be robust in explaining behavior in time and space. Among the many types of discrete choice models, the Multinomial logit (MNL) model is computationally the simplest and most practical (de Dios Ortúzar and Willumsen, 2011).

Destination choice models (DCMs) are a type of disaggregate trip distribution model that use discrete choice models, particularly logit models (Transportation Forecasting Resource, 2024). Adapting the definition of discrete choice models from de Dios Ortúzar and Willumsen (2011), DCMs operate on the fundamental principle that the probability of an individual choosing a given destination is a function of their socioeconomic characteristics and the relative attractiveness of the destination alternative.

Taking a simplified version of the example from Bernardin et al. (2009) for HBM trips, in Multinomial Logit (MNL) models of destination choice, the probability of a traveler residing in zone h, choosing a destination j for shopping/personal service (P_{hj}), is as follows:

$$P_{hj} = \frac{e^{W_{hj}}}{\sum_{j'} e^{W_{hj'}}} \tag{2.1}$$

where W_{hj} represents the traveler's utility, which is as follows:

$$W_{hj} = \ln \gamma S_j + \beta_c c_{hj} + \beta_{Hc} H_h c_{hj} \tag{2.2}$$

where:

- S_j: size of the destination with size parameter γ
- c_{hj}: Travel cost between home location h and destination j with cost sensitivity parameter β_c
- B_{Hc}: Parameter capturing interaction of traveler characteristic H with travel cost

During the 1970s, researchers began applying the MNL model, initially proposed by Suppes and Luce (1961) and rooted in utility maximization by Marschak (1974), to examine travelers' destination choices in academic studies (Ben-Akiva, 1974), among others (see Bernardin et al. (2009) for review). It has also been demonstrated that the MNL model is a generalized version of the gravity model (Daly, 1982).

Since the late 1970s, researchers have focused on the two key limitations of gravity and MNL models in spatial choice analyses. The first issue stems from the MNL model's Independence of Irrelevant Alternatives (IIA) property, which assumes uniform competition among all destination alternatives. According to the IIA assumption, the relative likelihood of selecting any two choices is unaffected by the presence of other alternatives in the choice set (Luce and Suppes, 1965). However, this contradicts the intuitive notion that all other factors being equal, nearby destinations are more likely to be substitutes and compete more intensely than distant destinations (Bernardin et al., 2009)

The stronger competition between nearby locations can be explained by several factors. One key reason is that it is often easier or cheaper to switch between locations that are close to each other, such as when changing destinations while traveling. Another reason is that, according to Tobler's First Law of Geography, all other things being equal, nearby places tend to be more alike than those farther apart, meaning that they have more in common. This similarity makes them more likely to compete directly with each other (Tobler, 1970). And since because of this, the nearby alternatives become correlated, the IIA assumption of the MNL models can result in unrealistic predictions (de Dios Ortúzar and Willumsen, 2011).

The second limitation is that gravity and MNL models do not account for the interdependency between multiple spatial choices made by the same individual. These models overlook the fact that a person's choice of workplace and other daily travel destinations are likely to be interrelated. It is reasonable to assume that people often select groups of nearby locations that can be visited together in a single efficient trip, thus minimizing their overall travel costs. This tendency to group destinations close to each other as part of a trip chain is an example of "economies of agglomeration" (Coe et al., 2007).

To address these limitations of gravity and MNL models, one solution is to enable the models to account for SC&AE. Two general modeling techniques have been proposed for this purpose. The first approach involves using models from the Generalized Extreme Value (GEV) family. This approach uses prior knowledge to group alternatives into different categories, allowing for varying levels of competition among them. For instance, when modeling work location choices, work locations can be grouped by region based on how close they are to each other. This grouping reflects the varying levels of competition between locations; nearby locations are seen as more competitive than distant locations. This enables GEV models to account for differences in spatial competition. This technique can also capture agglomeration effects; however, the model must be based on tours or activities, rather than individual trips. Following Shiftan (1998), this method of addressing agglomeration has been incorporated into a few tour-based and activity-based models (see Bernardin et al. (2009) for a review), although its adoption has been limited because of the high costs associated with data requirements, model development, and implementation (Ho and Hensher, 2016).

The second approach to addressing SC&AE involves using traditional MNL models but incorporating accessibility measures that include information about other potential destinations. By adding these accessibility measures, IIA assumption of the MNL model no longer applies. This method was first introduced by Fotheringham (1985) and has been adapted in many subsequent studies (see Bernardin et al. (2009) for a review) (Ho and Hensher, 2016).

2.2. Spatial Competition and Agglomeration Effects in Destination Choices

Spatial Competition and Agglomeration Effects (SC&AE) consider how the spatial distribution of opportunities across destination alternatives affects an individual traveler's destination choice. Spatial competition arises when opportunities in nearby zones decrease the attractiveness of a destination to a traveler. On the other hand, if the opportunities increase the destination's attractiveness to the traveler, it is called the agglomeration effect.

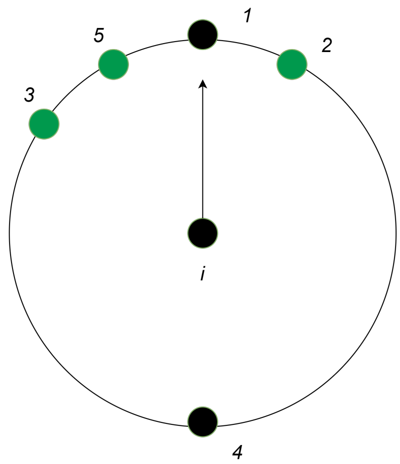


Figure 2.1: Agglomeration effect: Increase in attractiveness for destination zone 1 due to presence of other opportunities nearby

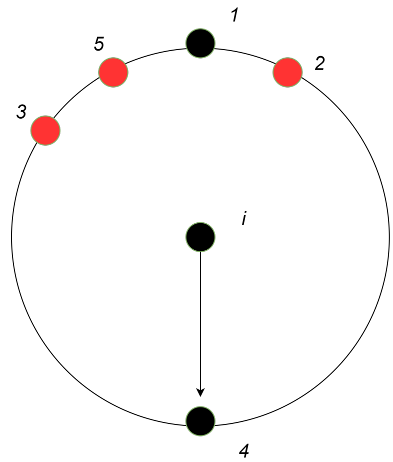


Figure 2.2: Spatial Competition: Decrease in attractiveness for destination zone 1 due to opportunities in nearby zones competing with it

To illustrate these effects, consider the example of home-based shopping trips adapted from Bhat et al. (1998). Consider the destination choice of an individual in zone i in the two spatial arrangements illustrated in Figures 2.1 and 2.2. All potential destinations (Zones 1 through 5) are equidistant from zone i and identical in all respects i.e., they are the same size and equally attractive. The traditional gravity model would predict the same trip volumes from zone i to each destination. At the disaggregate choice level, this means that the probability of choosing any of the Zones 1 through 5 as a destination is the same. However, the positioning of destination zones relative to one another can influence choice probabilities and, consequently, aggregate trip interchanges.

As shown in Fig. 2.1, the presence of several closely clustered shopping opportunities might lead individuals in zone i to perceive a greater variety or more opportunities for comparative shopping. This makes Zone 1 more attractive than the spatially isolated Zone 4, increasing its choice probability. This positive effect on a zone's attractiveness due to the proximity of other opportunities is known as the Agglomeration effect.

Alternatively, the same clustering of opportunities might lead to more competition for Zone 1, making it less attractive than spatially isolated Zone 4 in Figure 2.2. Individuals may also avoid congestion costs associated with zones with many nearby shopping locations. This negative effect on a zone's attractiveness due to the proximity of other opportunities is known as Spatial Competition, which reduces Zone 1's attractiveness and lowers its choice probability.

In reality, either competition or agglomeration effects may exist, and the appropriate effect can be determined from model estimation. One of the approaches introduced by Fotheringham (1985) of including SC&AE in MNL models is to use a Hansen-type accessibility index to include information about other destination alternatives. For the example illustrated in figures 2.2, and 2.1 and , for zone 1, the proximity of a destination zone to other shopping opportunities using a Hansen-type accessibility measure can be specified as:

$$A_1 = \ln \sum_{z=2}^{5} \frac{R_z}{c_{1z}} \tag{2.3}$$

where:
A_1: Accessibility index of destination zone 1
R_z: Sum of retail and service employment in zone z (a proxy for shopping opportunities in zone z)
c_{1z}: Travel cost between zones 1 and z

A generalized version of this equation is as follows:

$$A_j = \ln \sum_{z \neq j} \frac{R_z}{c_{jz}} \tag{2.4}$$

where:
A_j: Accessibility index of destination zone j
R_z: Sum of retail and service employment in zone z.
c_{jz}: Travel cost between destination zones j and z

Large values of A_j indicate more opportunities to shop close to zone j, whereas small values indicate that zone j is spatially isolated from other shopping opportunities. The accessibility index can be incorporated into the MNL form of DCMs by defining the utility function for destination zone j as a linear function of A_j. Consider the same utility specification described in equation 2.2, the specification can then be extended and further specified as a linear function of A_j and its parameter estimate β_A to account for SC&AE as follows:

$$P_{hj} = \frac{e^{W_{hj} + \beta_A A_j}}{\sum_{j'} e^{W_{hj'} + \beta_A A_{j'}}} \tag{2.5}$$

where β_A is the spatial structure parameter. If $\beta_A < 0$, zones close to other shopping opportunities have lower utility, indicating that the competition effects dominate. If $\beta_A > 0$, zones close to other shopping opportunities have a higher utility, indicating that agglomeration effects dominate. If $\beta_A = 0$, then there are no SC&AE effects or equally strong agglomeration and competition effects that cancel each other out. Using this accessibility measure, which includes information about alternative destinations, the MNL model's IIA assumption does not apply. Thus, including SC&AE in destination choice models with accessibility measures allows for spatial effect heterogeneity (Ho and Hensher, 2016).

A limitation of using accessibility indices focused on a single attraction is that they can only reveal the net effect of agglomeration and spatial competition, So even though these effects can occur simultaneously in case one effects dominates over the other, It isn't possible to determine whether the non-dominating effect exists or not. To address this, Bernardin et al. (2009) introduced destination choice models that separate these effects by using two different accessibility measures: one for complements (different types of attractions) and one for substitutes (same type of attractions), naming them Agglomeration and Competing Destination Choice (ACDC) models. These models were applied to non-work trips, such as shopping or personal business, and to work location choices by Ho and Hensher (2016).

This approach allows for a clearer understanding of the sources of competition and agglomeration effects, which can vary depending on trip purpose. For example, Bernardin et al. (2009) found that for non-work trips, spatial competition arises mainly from substitutes, while Ho and Hensher (2016) observed that for work trips, the same spatial competition comes from complements. A single accessibility measure can only show the overall effect and not the specific source. Although the numerical results might be the same as demonstrated by Bernardin et al. (2009) and (Ho and Hensher, 2016) , using two separate accessibility measures helps identify where these effects originate based on the type of trip. However, since this research focusses only on the impact of SC&AE on the temporal transferability of DCMS, this refined version of including SC&AE using two accessibility measures is not utilized in this research.

2.3. Role of Spatial Competition and Agglomeration in explaining destination choices across various trip purposes

SC&AE in DCMs have been a major focus of travel behavior research. As seen in table 2.1, various studies have explored these effects in diverse contexts, such as educational choices, workplace locations, and maintenance trip destinations in various regions globally. This sub-section focuses on research that has incorporated SC&AE into DCMs for trip purposes relevant to this study. Additionally, studies on educational location choices in the Netherlands that do not include SC&AE are reviewed to understand how other explanatory variables have been used to explain educational location choices in the context of the Dutch education system.

Table 2.1: Studies explaining destination choices across various trip purposes

Paper	**Area**	**Purpose**	**Incorporation of SCAE**	**DCM Type**	**Comments**	**Other Significant Parameters used**
Bhat et al. (1998)	Boston Metropolitan Area, USA	• Home based Work trips • Home based shopping trips	Only for Home based shopping trips using Hansen-type accessibility measure such that: Attraction factor: Log of retail and service employment in all zones Travel Impedance: Log of composite travel impedance from other zones. This composite travel impedance includes: In-vehicle travel time (IVTT), Out-vehicle travel time (OVTT), and travel cost and walk. The sum of the above term is divided by the total number of zones in the study area	MNL	A highly significant negative SCAE parameter indicating dominating competition effects.	• Log of travel impedance • Log of zonal size measure • Total employment (for work) • Retail + service employment(for shopping) • Interaction with travel impedance with: Gender, Age groups • Dummy variable to capture unique effects related to the central business district in the study area

Continued on next page

Table 2.1: Studies explaining destination choices across various trip purposes (Continued)

Bernardin et al. (2009)	Knoxville, Tennessee, USA	• Home-based Maintenance (HBM) • Home-based Others (HBO)	Uses two accessibility measures in the Agglomerating and Competing Destination Choice (ACDC) model: • Accessibility to employment in the same sector (Substitutes) • Accessibility to employment in different industries (Complements)	MNL	Sources of effects for HBM and HBO identified: • Agglomeration effects from complements • Competition effects from substitutes The net effect of SCAE is stronger for HBO trips compared to HBM trips. i.e., For HBO trips, people prefer closer locations to other destinations compared to HBM trips.	• For HBM trips. employments in Retail, Service • For HBO trips: Population,Enrollment, Retail, Service employments
Ho and Hensher (2016)	Sydney Greater Metropolitan Area (SGMA), Australia	Workplace Location Choice (WLC)	Adapted the ACDC model introduced by Bernardin et al. (2009): • Accessibility to jobs in the same industry (Substitutes) • Accessibility to jobs in different industries (Complements)	MNL	Sources of effects for WLC identified: • Agglomeration effects from substitutes • Competition effect from complements. when the above two parameters of are combined, its approximately the same as the net effect estimated using a single accessibility parameter	• Log of jobs relevant to the employee • Region specific constants • Logsum of mode and time of day choice • Region-specific constants

Continued on next page

Table 2.1: Studies explaining destination choices across various trip purposes (Continued)

Sá et al. (2004)	Netherlands	University Location choice for high school graduates	Uses a single accessibility called as the Centrality Index variable to capture the net effect. such that : Attraction factor: Total number of students in the universities Travel Impedance: Travel distance between universities	Production constrained gravity Model	Universities in densely populated areas (e.g., Randstad), the centrality index had a negative coefficient, indicating competition effects dominate and nearby institutions compete for students rather than benefit from agglomeration.	• Distance • Urbanization • Rent • Scope of university(Number of study programs offered)
de Boer and Blijie (2006)	Zwijndrecht, Netherlands	Primary education school choice	Not Applicable	MNL	Not Applicable	Models divided into 3 segments, households with: • Greater than modal income neigbourhood and western background • Lower than modal income and western background • Non-western background
van Welie et al. (2013)	4 major cities in Netherlands: • Amsterdam • Rotterdam • Utrecht • The Hague	Secondary education level: • HAVO • VWO • VMBO	Not Applicable	Ordinary Least Squares (OLS) regression	Not Applicable	There seems to be an interaction effect with the Migration background and Socioeconomic Status (SES) of student's residential neighborhood.

Building on the work of Fotheringham (1985), Bhat et al. (1998) examined home-based work and shopping trips in the Boston Metropolitan area (BMA). In this study, only for Home-based shopping trips, SC&AE is measured using a Hansen-type accessibility index (M_j), incorporating travel impedance factors (accounting for in-vehicle and out-of-vehicle times, and travel costs) and service and retail employment factors as an attraction factor for different zones.

$$M_j = \frac{1}{L} \sum_{l=1}^{L} \frac{\log R_l}{\log H_{jl}} \tag{2.6}$$

where:

- R_l: Represents the total retail and service employment in zone l, (a proxy for the shopping opportunities available in that zone).
- H_{jl}: composite travel impedance between zones j and l,
- L: Total number of zones within the BMA.

While this accessibility index could have been adapted for home-based work trips by considering total zonal employment, it was not applied to those trips in this study. For shopping trips, the results indicated a significant negative SC&AE parameter, reflecting dominant competition effects in shopping destination choices. The study also explored sociodemographic interactions with travel impedance, finding that older adults and women were more sensitive to travel impedance for work trips, while higher-income travelers were more willing to travel longer distances for work.

However, to address the limitation of the single accessibility measure of SC&AE highlighted in Section 2.1, Bernardin et al. (2009) proposed the Agglomerating and Competing Destination Choice (ACDC) model using two accessibility factors, which is applied to analyze home-based maintenance (HBM) and home-based other (HBO) trips. The two types of accessibility measures introduced are; one for substitutes (employment within similar sectors) and another for complements (employment across different sectors). Since this research does not explore this more refined version of including SC&AE in MNL DCMs, the equations are not elaborated further here. However, the results are quite insightful and this method can be explored in future research. The study finds that the competition effects, primarily from substitutes, are dominant for both trip types but stronger for HBO trips. This suggests that for discretionary activities, individuals tend to prefer locations that offer greater access to opportunities from different sectors more strongly, compared to maintenance trips. Ho and Hensher (2016) adapted the ACDC model to study workplace location choices (WLC) in the Sydney Greater Metropolitan Area (SGMA). The findings showed that, for work trips, competition effects were driven by complements, unlike the substitutes identified in non-work trips by Bernardin et al. (2009). Other important variables in the WLC destination choice model included the logarithm of relevant jobs (as an attraction factor), region-specific constants, and the logsum of mode and time-of-day choices, these choice models being a sub-component of a larger passenger travel demand model.

For educational purposes, a review of models explaining educational location choices in the Netherlands offers further insights. In their analysis of university location choices among high school graduates, Sá et al. (2004) incorporated SC&AE using a Hansen-type accessibility measure known as the 'Centrality Index' (c_j) within their production-constrained gravity model for trip distribution. The form is similar to the one used by Bhat et al. (1998) for shopping trips, albeit with a slight variation. This index incorporated the total number of universities as the attraction factor and the travel distance between universities as the travel impedance such that:

$$c_j = \sum_{m=1}^{n} \frac{P_m}{d_{mj}} \tag{2.7}$$

where:

- P_m: Attractiveness of university m, (Total number of students in the university) .
- d_{mj} : Road distance from university m to university j.

The findings indicated that universities in densely populated areas experience competition effects, with a negative centrality index suggesting that nearby institutions compete for students rather than benefit from agglomeration. Other significant variables in the study included are scope (number of programs offered), travel distance, urbanization level, and rental costs.

Regarding studies that did not consider SC&AE effects to explain education location choices, de Boer and Blijie (2006) focuses on primary school choice in Zwijndrecht, Netherlands. The study segments the population into three groups based on migration background and neighborhood income level to capture behavioral differences related to socioeconomic status and ethnicity. The motivation for this segmentation was to clarify how different groups prioritize factors like distance, school quality, and religious or alternative education, leading to more accurate predictions and tailored insights for school location planning. The study finds a general resistance to long-distance

travel for primary schools, but this resistance varies significantly depending on the student's background. For students from Western-origin households in higher-income neighborhoods, there is a strong sensitivity to travel barriers between home and school. These families are more willing to travel greater distances for schools with specific religious affiliations or alternative teaching methods (such as Dalton or Reformational schools). They also place a high value on school quality, preferring schools with better ratings and fewer disadvantaged students.
In contrast, students from Western-origin households in lower-income neighborhoods exhibit similar patterns regarding distance and school type, though to a lesser degree. For these families, traffic barriers and the number of students with disabilities are not significant factors in their school choice.
Students from non-Western backgrounds, predominantly Moroccan and Turkish, display high sensitivity to traffic barriers, similar to the higher income Western group. However, this group does not prioritize school characteristics such as religious affiliation or school quality in the same way as higher income Western group.

The study by van Welie et al. (2013) focuses on school location choice and upward mobility at the secondary education level among students in the Netherlands, specifically those in the secondary vocational track (VMBO) and the two academic tracks (HAVO and VWO). The analysis is based on BRON data [2], examining students living in the four major cities who were enrolled in the final grade of elementary school in 2008. To assess upward mobility, the study concentrated on pupils who began in the first year of the lower vocational track (VMBO) and were in their third year by the time of the research. Overall, selecting students residing in the four major Dutch cities (see table 2.1 for the names of cities) throughout the years 2008–2011. Although the study employs a different analytical approach, using Ordinary Least square (OLS) regression rather than choice or trip distribution models typically applied in transport research, it provides valuable insights into the factors influencing school location choice at the secondary level. The findings suggest an interaction effect between migration background and the socioeconomic status (SES) of the students' residential neighborhoods. A key factor influencing school choice is the sense of belonging, with students generally preferring schools with a larger proportion of peers from similar backgrounds.
Notably, native Dutch pupils are more likely to attend the nearest school when living in affluent neighborhoods, but they are more inclined to seek schools farther away when residing in lower SES areas, in contrast to migrant pupils from the same neighborhoods. It finds that socioeconomic indicators have a major influence on distance to school. Specifically, nonwestern pupils are more likely to select the nearest school when that school has a higher proportion of nonwestern students. Additionally, urbanicity, which reflects the level of human activity in a neighborhood based on the number of addresses, plays a role in students choosing a school farther away. In densely populated areas with more schools available, students have a wider range of options, which explains why they may not always choose the closest school. Similarly, the greater the number of schools within a 5 km radius, the more likely students are to consider schools beyond their immediate vicinity.

2.4. Factors affecting temporal transferability of choice models

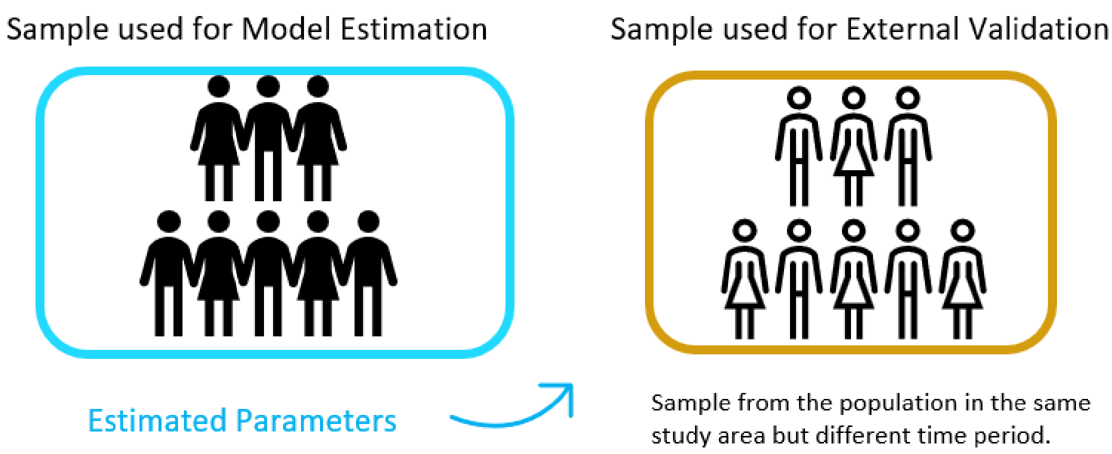


Figure 2.3: Assessing Model's Temporal Transferability via External Validation

[2]Basis Register Onderwijs (BRON): Dutch official Educational database

In travel behavior, transferability can be defined as "the ability of a model developed in one context to explain behavior in another, assuming the underlying theory is equally applicable in both contexts" (Fox et al., 2014). Thus, temporal transferability is the model's ability to maintain its accuracy and reliability over a forecasting horizon without the need for extensive recalibration.

Assessing temporal transferability helps to determine how well a model can adapt to changes over time, providing confidence in its predictions for transport planning and its ability as a tool to aid well-informed decision-making. This concept is particularly relevant in transportation planning, where models are used to predict future travel behaviors based on past data (Fox et al. (2014); Parady et al. (2021)).

Temporal transferability is a type of external validation. External validation tests the generalizability and accuracy of a model beyond the conditions or population for which it was originally developed. It involves applying the model to a new, independent dataset unused during the model's development. This process helps verify whether the model can maintain its prediction accuracy when subjected to new data. When this new data is from a different time period, external validation assesses the model's temporal transferability, providing insights into its reliability over that forecasting horizon (Parady et al., 2021).

Figure 3 shows how external validation is used to assess the temporal transferability of the model. First, a sample from the population is used for model estimation. Once the parameters are estimated using the base year sample, these same parameters are then applied to predict choices observed in the population collected from the same study area in a subsequent time period.

Table 2.2 and 2.3 outline the factors that influence the temporal transferability of transport choice models. There is limited research focusing exclusively on the temporal transferability of destination choice models. Hence, much of the existing literature reviewed here pertains to mode choice and mode-destination choice models. It is well recognized that what may initially appear to be separate decisions, selecting a destination and choosing a mode of transportation are joint decisions involving combinations of mode and destination options (Castiglione et al., 2014). Owing to the interdependence of these two aspects of travel decisions, In ABMs, destination choice, and mode choice are linked sequentially because the selection of a destination influences the choice of transportation mode (Castiglione et al. (2014); Clifton et al. (2016); Jonnalagadda et al. (2001)). The destination choice model determines where an individual chooses to travel, considering factors such as trip purpose, distance, and accessibility, whereas the mode choice model predicts the transportation mode based on the chosen destination and other relevant factors (Clifton et al. (2016); Lekshmi et al. (2016)).

Table 2.2: Factors Affecting Temporal Transferability of Mode and Mode-Destination Choice Models

Factors	Effect	Note	Study	Choice Model and Purpose	Area	Time Frame
Regional characteristics such as Infrastructure, Population growth, land use etc	Significant change in the forecasting year compared to the based year can negatively impact the transferability of the choice		Sanko and Morikawa (2010)	Mode Choice for: • Commuter trip • Business Trip	Chukyo metropolitan area, Japan	20 years (1971,1991)
Past travel behavior (Inertia)	Higher the inertia, better the temporal transferability	Changes in travel behavior are: • Path dependent: Depends on the direction of change of Level of service • Anti-symmetric: The degree of behavior change to LoS change varies by transport mode.	Sanko and Morikawa (2010)			
Forecasting Horizon	Transferability declines as horizon increases	No consistent pattern seen in the results. Therefore no conclusive evidence.	Fox et al. (2014)	Commuter Mode-destination choice	Greater Toronto and Hamiton Area	20 years (1986-2006); tested across 1986, 1996, 2001, and 2006
Overfitting	Higher tendency to overfit reduces temporal transferability		Fox et al. (2014), Parady et al. (2021)			

Table 2.3: Factors Affecting Temporal Transferability of Mode and Mode-Destination Choice Models (Continued)

Factors	Effect	Note	Study	Choice Model and Purpose	Area	Time Frame
Model specification	Improved specification improves transferability	Improvement in specification by adding behavioral parameters (socioeconomic) rather than constants improves temporal transferability.	Fox et al. (2014)			
			Parody (1977)	Commute, Mode choice	Univ. of Amherst, Mass., U.S.	4 phases: 1972 Autumn, '73 Spring, '73 Autumn, '74 Spring
Type of parameters: • Level-of-service parameters (LoS)	LoS parameters are more transferable/stable, therefore potentially improve transferability.		Fox et al. (2014)			
• Cost parameters	Cost parameters less transferable, possibly limiting improvements in transferability		Fox et al. (2014)			
IIA assumption	Addressing the IIA assumption of MNL logit models could improve transferability.		Train (1978)	Commute, Mode choice	San Francisco, U.S.	Pre BART[6] sample (before 1973) compared to post-BART sample (before 1976)

As shown in Table 2.2 and 2.3, several factors affect the temporal transferability of mode and mode-destination choice models. One key factor is regional characteristics and past travel behavior also termed 'inertia' (Sanko and Morikawa, 2010). Significant changes in regional characteristics during the transfer year can substantially alter model parameters compared to the base year, potentially undermining transferability by introducing behaviors not accounted for in the original model. Fox et al. (2014) finds that improving the model specification enhances transferability, which aligns with empirical findings from previous studies on mode choice models (see Fox and Hess (2010) and Fox et al. (2014) for a detailed review). This improvement stems from including behavioral parameters such as socioeconomic variables in the model specification, which reduces the reliance on constants to explain behavior in choice models. Here, Fox et al. (2014) finds that the constants are the least stable model parameters, and as the influence of these constants diminishes with the addition of behavioral parameters, the transferability of the model improves. One of the previous studies by Parody (1977), conducted a before and after methodology to understand to validate how well disaggregate logit models predict the changes. In the study, the research focused on how mode choice among travelers changes by introducing a free bus service and then subsequently increasing parking fees and implementing stricter parking regulations. Conducted in the University of Massachusetts at Amherst, it focused on travel behavior changes and compared the model predictions by a disaggregate mode choice logit model across four key periods: Fall 1972 (baseline, before any transportation changes), Spring 1973 (after implementing an expanded free bus service), Fall 1973 (following increased parking fees and stricter parking regulations), and Spring 1974 (during the energy crisis with increased gasoline prices).

The study finds that Disaggregate modal-choice models, particularly those incorporating detailed socioeconomic (such as gender, and occupational status of the traveler) and transportation service variables (Frequency of service, Walk time to the bus stop), demonstrated much better predictive accuracy in forecasting shifts in travel modes resulting from transportation system changes. It also observed that across the 4 periods, travel time had the most stable coeffiecient.
Although Fox et al. (2014) initially hypothesized that transferability would decrease as the forecasting horizon lengthens, the study did not find a consistent pattern across different years to conclusively support this hypothesis. Nevertheless, it is reasonable to expect that over longer forecasting horizons, substantial changes in regional characteristics such as demographic shifts, land use, effects of policy implementations, and technological developments such as new transport services and improved accessibility would lead to changes in travel behavior. In terms of parameter stability, Fox et al. (2014) also found that level-of-service parameters are generally more transferable than cost parameters, which may be related to the past travel behavior or 'inertia' factors discussed by Sanko and Morikawa (2010).
Using a similar before and after approach as done by Parody (1977), Train (1978) examined the predictive accuracy of a mode choice model focusing on the Bay Area Rapid Transit (BART) system opened in San Francisco. By analyzing data from workers surveyed both before and after BART's launch, the study validated the model using two approaches: comparing post-BART actual mode shares to those predicted by the pre-BART model and assessing the stability of parameters across pre- and post-BART models. The findings indicate that the model tends to overestimate transit use, particularly for BART with walk access. To assess whether the IIA assumption is a potential reason for the over-prediction of transit usage, the study tested alternative models that did not depend on IIA, such as the Maximum and Log-sum models. These non-IIA models also overestimated transit usage, indicating that the IIA assumption failure was not the main source of error. Ultimately, the study found that although the IIA assumption might slightly contribute to overprediction, other issues, like unique BART-specific attributes and inaccurate walk time data, were the primary contributors to the overprediction.
Finally, multiple studies have cautioned against the risk of overfitting models to current data, as this can negatively affect forecasting accuracy. This risk of overfitting means it is dangerous for modelers to believe that an improvement in model fit automatically translates into better predictive power across different forecasting horizons, which is not necessarily the case (Fox and Hess (2010); Fox et al. (2014); Parady et al. (2021)).

2.5. Discussion

Parady et al. (2021) discusses in detail why evaluating the performance of transport models based solely on improvements in goodness-of-fit criteria undermines their credibility as predictive tools and effective decision-making instruments for transport planning. The primary concern is the risk of overfitting. A model that can accurately explain current travel behavior does not mean that it will be equally effective at predicting future travel patterns. This is why improvements in model specification should not be justified or rejected solely based on their contribution to explanatory power.

Overfitting occurs when a model is heavily dependent on the context to explain travel behavior. The more overfitting the model is, the more it will lose its explanatory power with variation in the context, making the model less reliable. This means that the model explains travel behavior that is highly specific to the current context, rather than the fundamental drivers of travel decisions. Consequently, even if new theories enable the model to explain current behavior more precisely or estimate parameters more accurately, they may undermine the model's predictive ability over the desired forecasting horizon. These improvements could negatively impact the model's temporal transferability, meaning that it may not perform well in future scenarios. Thus, relying on a model's explanatory power as a measure of its predictive capability and using it to make policy recommendations for future travel behavior is both dangerous and irresponsible.
However, this does not mean that theories that enhance a model's ability to explain behavior are redundant. Instead, the above arguments are a caution against assuming or assessing their contribution to a model's predictive power based on tests designed to evaluate their contribution to explanatory power. It is crucial to determine whether the model is intended as an explanatory tool or a predictive one. When improving model specifications, it is important to understand the purpose of these enhancements. Moreover, the contribution of theories to a model's performance should not be viewed in binary terms, either enhancing explanatory or predictive power, but rather on a spectrum. If we consider this spectrum only in two aspects; Explanatory and Predictive power, incorporating theories and improving model specifications can:

1. Improve explanatory power but reduce transferability (by capturing highly contextual behavior, tending to

overfit).

2. Enhance transferability with little impact on explanatory power (by capturing fundamental drivers of behavior that are applicable across different contexts).
3. Significantly improve both explanatory power and transferability.
4. Limited impact on both aspects.

Therefore, when refining model specifications with a mix of theories and parameters, transport modelers need to assess where the needs of the user or stakeholder in transport planning stand on this spectrum and decide accordingly.

Continuing with this argument, it is inappropriate to judge the significance of a theory or improvement based solely on one criterion, such as its contribution to explanatory power, using Goodness of fit statistics as observed as a common trend in current transport research by Parady et al. (2021). Doing so creates a myopic perspective in our already limited understanding of human travel decision-making. We need a multifaceted evaluation approach because these choice models attempt to represent a highly complex travel decision-making process. Justifying or excluding a theory based on only one aspect of model performance is overly restrictive because it fails to account for the multidimensional nature of the travel decision-making process that it is designed to represent. When we assess only one aspect, we implicitly assume that only this aspect matters for the multidimensional system the model is supposed to represent. Travel behavior is influenced by multiple factors, both of which are known and yet to be discovered. Hence, focusing on only one aspect could lead to misguided conclusions regarding the importance of the theory in our quest to understand travel behavior better. To capture this complexity requires a holistic approach. Assessing the theory's impact on temporal transferability in addition to the explanatory power is just one of the many aspects of model performance and, hence, a step towards a more holistic approach.

Returning to the theory of SC&AE, it has been proven in the transport research literature that these effects significantly explain various types of destination choices, such as university destinations in the Netherlands (Sá et al., 2004), work locations in Sydney (Ho and Hensher, 2016), home-based maintenance trips (shopping, personal errands), and home-based other trips in Knoxville and Tennessee (Bernardin et al., 2009). Developed in 1985 by Fotheringham (1985), this theory has stood the test of time, with these recent studies demonstrating its significant explanatory power for travelers' destination choices in different regions across the globe. Given its long-lasting relevance and application in various contexts, it seems that the theory captures a fundamental human behavior that drives how travelers choose destinations. Previous research on temporal transferability indicates that explaining travelers' choices using behavioral parameters improves the temporal transferability of the model (See section 2.4).

Another advantage of incorporating SC&AE, as noted in the previous section 2.1, is that it relieves the MNL form of destination choice models from their popular IIA assumption. This IIA assumption can lead to unrealistic predictions in cases where alternatives are similar or share unobserved attributes i.e., correlated alternatives (de Dios Ortúzar and Willumsen, 2011). Thus, SC&AE enables retaining the simplicity of MNL models, which makes them computationally less intensive and easier to use for large frameworks, such as activity-based modeling or trip-based models. This simplicity is especially advantageous as DCMs are a subcomponent for trip distribution in these large frameworks.

All these arguments highlight SC&AE's high potential to positively improve the temporal transferability of destination choice models. Yet, no research has investigated the impact of its inclusion on the temporal transferability of destination choice models. By providing quantitative evidence of this theory's contribution to such a crucial and reliable feature of transport models, this study aims to justify its inclusion comprehensively.

3
Data

This chapter provides an overview of the data sources and assumptions regarding travel time used in this study to develop destination choice models in Chapter 4 for the Metropolitan Region of Amsterdam (MRA). The datasets used include the Dutch National Travel Survey, employment data, educational enrollment, and travel time matrix.

This chapter also discusses how the centroid for the destination zones is determined, the assumptions made for estimating intrazonal travel times, and how the mode-specific travel time matrix for each trip purpose was selected.

3.1. Dutch National Travel Survey Data (ODiN)

Onderweg in Nederland[7](ODiN) is a travel survey that tracks the travel behavior of the Dutch population. The participants are required to record their daily travel details, including destinations, purposes, modes of transport, and travel duration, for one specific day each year. In addition, they provide information on general personal and household characteristics and details about driving licenses and available modes of transport (DANS, 2024). Appendix B lists the potential variables from this survey that can be added to the utility specification as traveler characteristics for various trip purposes. The trips are filtered to include only those arriving in the municipalities included in the MRA. The MRA region consists of 30 municipalities. The arrival points of the trips are available at the municipality level and the 4-digit postal code (PC4) level. The rest of the data below are processed at the PC4 level to match the minimum resolution of the travel survey (see Section 4.3).

(skip this paragraph, if you already have read the scope in Section 1.2)
ODiN is a continuation of the "Onderzoek Verplaatsingen in Nederland" (OViN) survey, which was carried out by CBS from 2010 to 2017. The methodology of ODiN 2018 differs significantly from that of the earlier OViN survey, creating a methodological break between the two. As a result, the findings from ODiN cannot be directly compared with those from OViN analyses (DANS, 2024). ODiN surveys only individuals aged six and older, excluding younger children, which reduces the target population by over 1 million. It uses an internet-only (cawi) data collection method, unlike OViN's mixed-mode approach. ODiN also integrates all domestic and international vacation trips as regular trips, whereas OViN handles these trips separately. In addition, ODiN determines the primary mode of transport based on distance, unlike OViN's priority-based system. Finally, ODiN utilizes more register-based data called Basisregistratie Personen (BRP), reducing the need for extensive survey questions (Statistics Netherlands (CBS), 2019).
Although adjustments could improve the comparability between ODiN and OViN, some differences would likely remain due to inherent methodological changes. Given the limited timeframe, making such modifications is beyond the scope of this study. Therefore, considering that the earliest available ODiN data is from 2018 and the most recent from 2022, a short-term forecasting horizon of five years (2018-2022) is selected for this analysis.

3.2. Employment data

This data represents the number of jobs in the MRA, with the highest resolution available at the municipal level and the lowest available at the PC4 level.It is provided by Research and Statistics, Amsterdam (O&S). The data structure

[7]English translation: The Survey on the Road in the Netherlands

is detailed in Table B.5 in Appendix B. This includes the total number of jobs in each of the 30 municipalities. This information is used as a zonal size measure in the utility function for work location choice. Additionally, the data categorizes jobs into wholesale, retail, and other services, which are summed and then used as a zonal size measure for HBM trips as per the utility functions defined for the trip purpose in the design phase of the methodology.

3.3. Education Enrollment data

This dataset provided by Dutch Ministry of Education, Culture, and Science (DUO) contains the locations of educational institutions across the Netherlands and the number of enrolled students in them. It covers the primary, secondary, vocational, and higher education levels corresponding to the Dutch education system (See Appendix E). Because the number of students is provided at the institution level, postal codes are determined with the help of location information (see Sections 4.3 and 4.3.3), enabling the number of enrollments to be aggregated at the PC4 level. This information is then used as an attraction factor for educational trips in the MRA. The data structure is presented in Appendix B and in Table B.6.

3.4. Travel time matrix

The travel time matrix between all the PC4 zones in the Netherlands for various private transport modes (walking, bicycle, and car) and public transport modes is calculated using the GeoDMS software developed by ObjectVision. The software is continuously developing and actively used by organizations such as the Netherlands Environmental Assessment Agency, Joint Research Centre of the European Commission, Vrije Universiteit Amsterdam, and various Dutch municipalities (Object Vision, 2024). Appendix C details the input parameters used for calculating the travel time matrix for various modes in the GeoDMS software. The source data for this calculation are as follows:

Road Network Data

OpenStreetMap provides detailed data for each road and path segment, including information on road type, name, and whether a street is one-way (ObjectVision, 2023). Network data from parts of neighboring countries such as Belgium and parts of Germany (Niedersachsen and Nordrhein-Westfalen) are also included, as some of the shortest paths between PC4 zones in the Netherlands pass through these areas.

Geolocation of Buildings and Addresses (BAG)

The BAG (Basisregistratie Adressen en Gebouwen)[8]is the registry containing geolocation information about all buildings and addresses across municipalities in the Netherlands. This data is used to determine the centroids of the PC4 zones.

Public transport Data

The General Transit Feed Specification (GTFS), Google's publicly available feeds, is used to build routes between the origin and destinations (these points are defined using BAG data) using the PT Mode. It comprises all forms of public transportation, including trains, metros, trams, buses, and ferries. It provides details on all stops and departure and arrival schedules and covers all designated transport routes.

3.4.1. Determining PC4 Centroids

In accessibility studies, when calculating travel impedance (such as travel time), the geometric center of a region is often used as its representative point. This geometric center is calculated based on the geographical shape or boundaries of the area. However, the geometric center may not always accurately reflect where the population intends to travel or where points of interest in a region are concentrated. It may fall in an isolated or uninhabited part of the region. To address this, instead of using the geometric center, the address-weighted center of gravity approach considers the locations of all the addresses within the area. This approach provides a more accurate representation of where people and opportunities are concentrated, thus better reflecting actual travel patterns.

Figure 3.1 illustrates the differences between these two approaches, highlighting how the address-weighted center of gravity based on BAG addresses differs from the geometric center based solely on geographical boundaries. In many cases, these two points may be close; however, the difference can be substantial in regions where buildings

[8]English translation: Key Register of Addresses and Buildings

are clustered toward a specific part of the area. The geometric center of each zone is represented by a grey triangle, calculated based on the geographical boundaries of the area (blue lines). However, the black dots indicate the address-weighted center of gravity, which accounts for the region's distribution of all buildings and addresses, as derived from the BAG (Basisregistratie Adressen en Gebouwen) dataset. The positions of the black dots differ from those of the gray triangles, demonstrating the influence of building density on determining a more population-relevant center point. In areas where buildings are clustered, the address-weighted center of gravity is noticeably shifted compared to the geometric center. (ObjectVision, 2023).

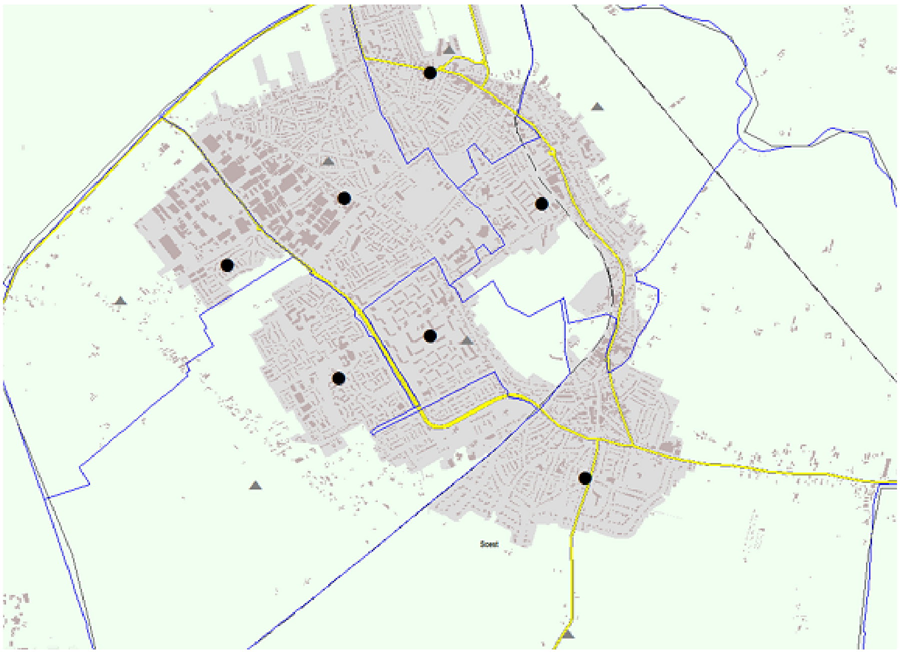


Figure 3.1: Difference in the geometric center (grey triangle) and the center of gravity based on buildings (black dots) for a region (ObjectVision, 2023)

3.4.2. Assumptions for Intrazonal and trip-specific travel times

Because the travel time matrix is only available at the PC4 level, travel times for intrazonal trips, i.e., trips occurring within the same PC4 zone, are not directly provided and must be assumed. Following a standard rule of thumb, the travel time for intrazonal trips is assumed to be half the travel time to the nearest neighboring zone (Cats, 2022).

Travel time matrices for different modes, such as car, public transport (PT), bike, and walking, are calculated using GeoDMS. Ideally, the travel times for all modes should be incorporated into the travel time matrix to account for the availability of various transportation options. One approach to achieve this is to examine the frequency of each mode used in the travel survey, grouping them into four main categories: car, PT, bike, and walk, and then creating a combined travel time matrix. This matrix would reflect the weighted average travel times based on mode usage frequencies. Table 3.1 illustrates how various modes used for HBM trips (fig 3.2) are grouped.

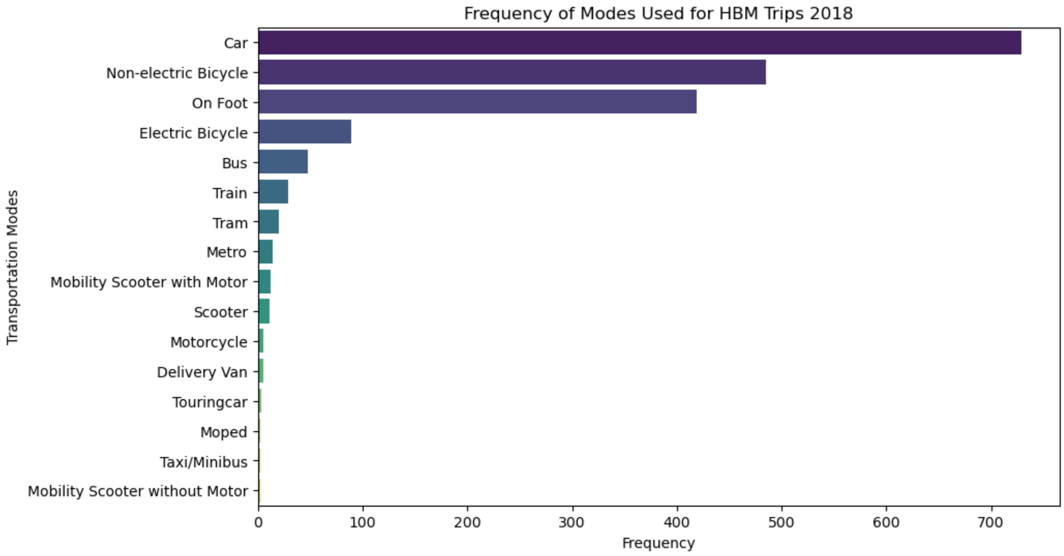


Figure 3.2: Frequency of Modes Used for HBM trips in ODiN 2018

Table 3.1: Grouping of Modes used for HBM trips, ODiN 2018

Car	PT	Bike	Walk
Car	Train	Non-electric Bicycle	Walk
Touring car	Bus	Electric Bicycle	
Delivery Van	Tram	Mobility Scooter with Motor	
Taxi/Minibus	Metro	Mobility Scooter without Motor	
Motorcycle			
Moped			
Scooter			

As per this grouping, the following percentages of the four categories are illustrated in fig 3.3

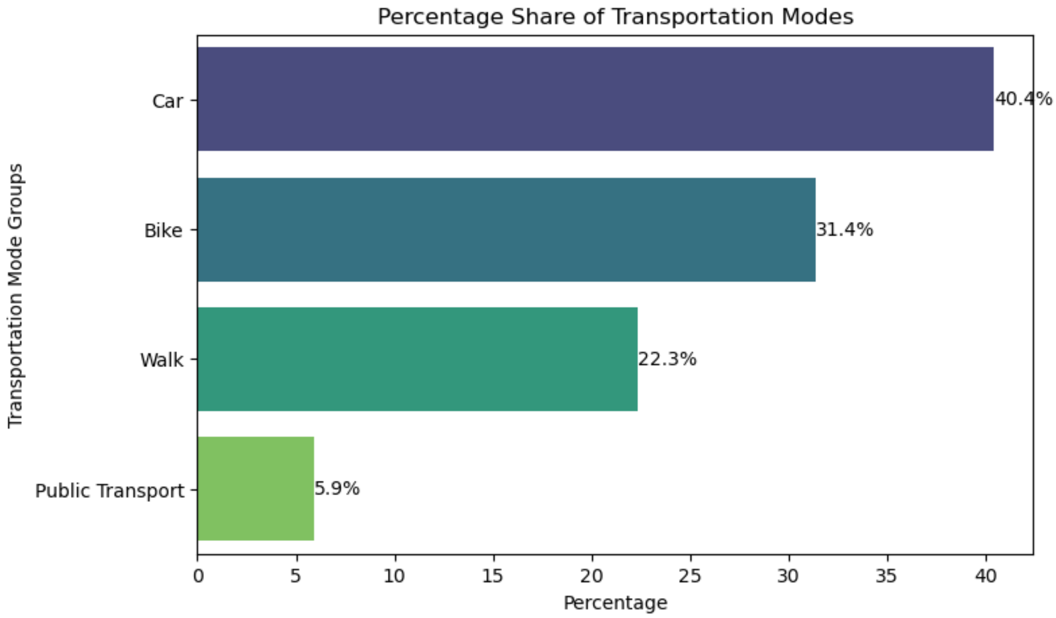


Figure 3.3: Percentage share of modes used for HBM trips, ODiN 2018

However, applying weighted mode share percentages to incorporate all modes results in unrealistically high travel times between zones, particularly for distant zones. This is likely due to the significant contribution of walking times, which constitute 22% of the total, inflating travel times considerably. Additionally, combining travel times for all modes leads to exaggerated travel durations. In some cases, the calculated travel time between zones can reach up to 10 hours.

A simpler approach was adopted to address this issue. Instead of including all modes, this study uses the travel time for the mode with the highest frequency for each trip type. Under this assumption, the mode-specific travel time matrices listed in Table 3.2 are selected for each trip category.

Table 3.2: Mode with highest frequency per trip type

Year	HBM	Work	Primary Education	Secondary and above trips
2018 2022	Car	Car	Bicycle	Bicycle

The frequency count per mode for each year per trip type is listed in their respective sections in Appendix F. HBM (fig F.8), Work (fig F.5), Primary Education (fig F.11) and Secondary and above (fig F.13)

4
Methodology

This chapter elaborates on the research methodology used in this thesis. The chapter begins with an overview (Section 4.1). Here, it gives an overall gist of the various phases of the methodology used in this thesis to answer the research question and how they are connected.

After this overview, this chapter is structured into several key sections. Section 4.2 discusses the selection of variables based on the findings from the previous chapter Literature Review. Variables are chosen based on trip purpose, available data, and factors affecting temporal transferability. The variables include zonal size measures (such as employment and education enrollment), travel impedance, sociodemographic interactions with impedance, and SC&AE.

Section 4.3 details the steps taken to prepare the datasets for the model estimation. This includes processing the ODiN travel survey data employment data, and education enrollment data for 2018 and 2022. The ODiN data is filtered to include relevant trips within the MRA for work, home-based maintenance (HBM), and education purposes. Issues such as duplicate trips, missing values, and aggregation to the PC4 zone level are addressed. Employment and education data are aggregated at the PC4 level and merged to create a master dataset for model estimation.

In Section 4.6, utility functions for different trip purposes are formulated. For work and HBM trips, the utility includes relevant zonal size measures (employment), travel time, and interactions between sociodemographic variables and travel time. For education trips, utility functions are specified separately for primary and secondary & higher education, using enrollment numbers and relevant sociodemographic interactions. SC&AE is included using a Hansen-type accessibility index to capture the influence of the spatial distribution of opportunities.

Addressing the need for computational efficiency and realistic behavioral representation in forming a choice set, Section 4.7 elaborates on a variant of Stratified Importance Sampling (SIS) used in this thesis, which considers both proximity and oppportunies present in the destinations.

Section 4.8 discusses the steps for assessing the impact of SC&AE on the temporal transferability of the destination choice model across trip purposes. Further, the selection of the performance indicators based on which the assessment of impact will be performed is discussed, along with how the same steps will be applied to Full (no sampling), random sampling along with SIS to understand the impact of SC&AE on temporal transferability varies with sampling methods.

4.1. Overview

Figure 4.1 presents an overview of the methodology used to address the research questions outlined in Section 1.3. The methodology is organized into three distinct phases: Exploration, Design, and Validation. Each phase sequentially addresses the sub-questions (SQ) that collectively contribute to answering the main research question (RQ).

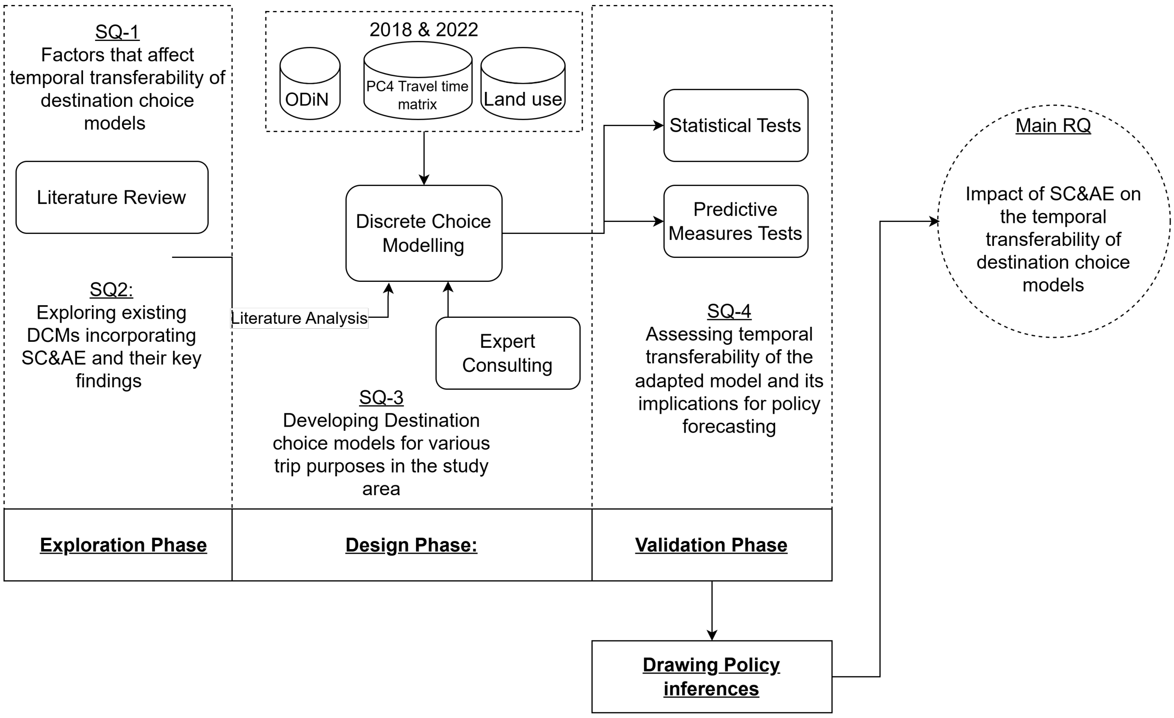


Figure 4.1: Methodology overview

The **Exploration Phase** begins with a literature review to identify the factors influencing the temporal transferability of mode and destination-mode choice models (Section 2.4) and to examine how previous studies have accounted for Spatial Competition and Agglomeration Effects (SC&AE) across different trip purposes. In addition to SC&AE, the review examines other variables included in these studies, exploring how they were specified to explain travelers' destination choices (Section 2.3). The next step is to determine which of these variables, along with SC&AE, are available within the four datasets mentioned in Chapter 3 and how they can be incorporated to explain destination choices across trip purposes. Hence, the structure and available information for the four datasets are reviewed. Together, the literature and data review forms the basis for variable selection (Section 4.2) and data processing. This phase establishes a robust theoretical foundation for designing Destination Choice Models (DCMs) that account for SC&AE in the study area across the three trip purposes.

The **Design Phase** builds on the findings of the exploration phase to develop utility specifications for Multinomial Logit (MNL) destination choice models designed for each trip purpose.

During the **Validation Phase**, the parameters of the trip-specific MNL models are estimated using the open-source Python package, Pandas Biogeme. First, the parameters are estimated separately for the 2018 and 2022 datasets. Then the 2018 estimated parameters are used on the 2022 dataset to assess the impact of SC&AE on temporal transferability. Four performances across three categories of indicators are selected to evaluate the models. These indicators are chosen to complement one another, addressing distinct aspects of discrete choice model performance. This approach provides a comprehensive evaluation of SC&AE's impact on temporal transferability across various trip purposes.

Finally, the thesis concludes by **drawing policy inferences** from insights gained during the validation phase. This involves reflecting on SC&AE's role in increasing the effectiveness of destination choice models and discussing the broader implications for developing transport models as a reliable tool for transport policy planning.

4.2. Selection of Explanatory Variables

To assess the impact of SC&AE on the temporal transferability of destination choice models, a utility specification for each trip purpose needs to be formulated. Hence, we must select the variables to include in the specification. The selection of explanatory variables differs with the trip purpose. The selection also depends on available data and factors affecting temporal transferability. Because the highest common resolution across all four datasets mentioned in the Data Section is at the PC4 level, PC4 zones in the MRA region are considered as destination alternatives.

For including explanatory variables, in addition to the SC&AE parameter, three sets of explanatory variables are considered to be included in the MNL choice models: (a) Zonal size measures, (b) Travel impedance, (c) Interaction of sociodemographic variables with impedance, and (d) SC&AE

Zonal size measures

Zonal size measures vary according to the trip purpose. Typically, for work, total employment is considered; hence, total employment in the PC4 destination zone is considered in this thesis. For HBM, employment across relevant sectors is considered. These include six sectors: Wholesale & Retail, Financial Institutions, Utilities, Government, Health & welfare, and Other services. Lastly, for education trips, total enrollment at the relevant education level is considered. The size term is always included in the utility function in a logarithmic form. This ensures that the probability of selecting a destination is directly proportional to the number of opportunities available in the destination zone (Bernardin et al., 2018).

Travel Impedance

For travel impedance across all trip purposes, typically, the logarithmic transformation of travel time has been used for explaining destination choices in the studies reviewed in Section 2.3. Hence, the logarithm of travel time between PC4 zones is considered here as well. This is equivalent to the power function of the impedance function of the gravity model (Daly, 1982). Travelers often are more sensitive to changes in travel time when the time is short compared to when it is long. As seen in figure 4.2, the power function captures this diminishing sensitivity.

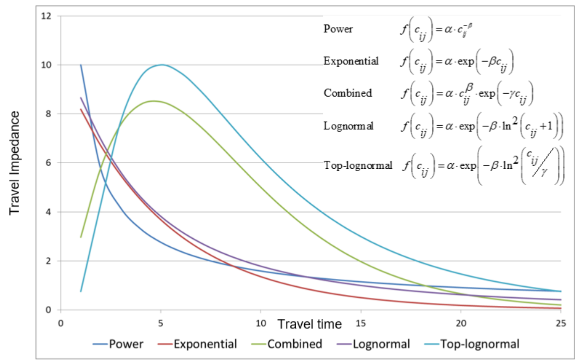


Figure 4.2: Different types of travel impedance functions (Cats, 2022)

Interaction of sociodemographic variables with travel impedance

As reviewed in Section 2.4, adding the sociodemographic characteristics of travelers in the specification improves the temporal transferability of destination choice models. The interaction of sociodemographic variables with impedance varies with the trip purpose. These interactions are included based on findings from previous studies related to the three trips within the scope of this study. Bhat et al. (1998) finds that older adults and women are more sensitive to travel impedance for work trips, while higher-income travelers are more willing to travel longer distances for work. Hence, the interactions of gender, age, and disposable household income level of travelers with travel time are included for work and HBM trips.

For education trips, previous research focusing on Dutch education, such as de Boer and Blijie (2006) (primary education) and van Welie et al. (2013) (One level after secondary education), finds highly significant interaction of socioeconomic status and migration groups of students with travel impedance. Hence, the interactions of migration and disposable household income levels of travelers with travel time are included in education trips.

Spatial Competition and Agglomeration Effects
Finally, to evaluate the impact of SC&AE on the temporal transferability of destination choice models and how it varies by trip purpose, these effects are included using a single Hansen-type accessibility index, as shown in equation 2.4.

4.3. Data Processing

This section details the data processing steps necessary for model estimation using Pandas Biogeme. It covers the processing of ODiN travel survey data (subsection 4.3.1), employment data (subsection 4.3.2), and education enrollment data (subsection 4.3.3).

For the ODiN data, trips from 2018 and 2022 are filtered to focus on trips to the 30 municipalities within the MRA, specifically for Work, Home-based Maintenance, and Education purposes using SQL. Subsequently, the data is processed in Python to address issues such as duplicate trips, missing values, and aggregation to the PC4 zone level.

The employment and education data for the 364 common PC4 zones from 2018 and 2022 are combined with the processed ODiN data to create a master dataset. This master dataset is then used for DCM estimation for the trip purposes in Pandas Biogeme.

4.3.1. ODiN data processing

This subsection outlines the step-by-step data-processing workflow for the ODiN survey data for 2018 and 2022. As noted earlier, The ODiN data is filtered only to include arrivals in the 30 municipalities inside the MRA region as shown in Appendix D. To do so, the below filters in Table 4.1 are applied for the three trip purposes. i.e., Work, Home-based Maintenance and Education to the ODiN data stored as an SQL database in TNO's repository.

Table 4.1: SQL filters for processing ODiN Data

Variable explanation	**Variables**	**Type of trips**				**Notes**
			Home based			
		Work	Shopping	Service	Education	
	To Include					
New move	verpl	1: Yes	1: Yes	1: Yes	1: Yes	
Motive Classification	kmotiefv	1: To and from work	4: Shopping/grocery shopping	3: Services/Personal care	5: Education/course	
Purpose	doel	2: To Work	≠ 1			1: Going Home
Departed from	vertloc	—	1: Home		—	
Arrival Gemeente (Municpality)	aankgem	358,34,362,363,375,376,377,384, 385,1942, 392,394,396,397,402,406, 415,417,995,431, 437,439,450,451, 453,852,1696,880,479,473				These are the numbers of the 30 gemeente in the MRA region (see Appendix D)
	To exclude					

Continued on next page

Table 4.1: SQL filters for processing ODiN Data (Continued)

Number of passenger cars in the household	hhauto	10, Null	—	10: Unknown
Traveler's paid work status	betwerk	4, Null	—	4: Unknown
Traveler's car license status	oprijbewijsau	2, Null	—	2: Unknown
Student OV chip card status	ovstkaart	—	3, Null	3: Unknown
Arrival PC4 zone code	aankpc	0, Null		
Departure PC4 zone code	vertpc	0, Null		
Standardized household disposable income (Decile Groups)	hhgestinkg	11, Null		11: Unknown
Migration Background	herkomst	4, Null		4: Unknown
Number of regular trips in Netherlands	aantvpl	1		
Highest completed education	opleiding	5, 6		5: Other training, 6: Unknown
Number of movements by OP	weggeweest	0: No		

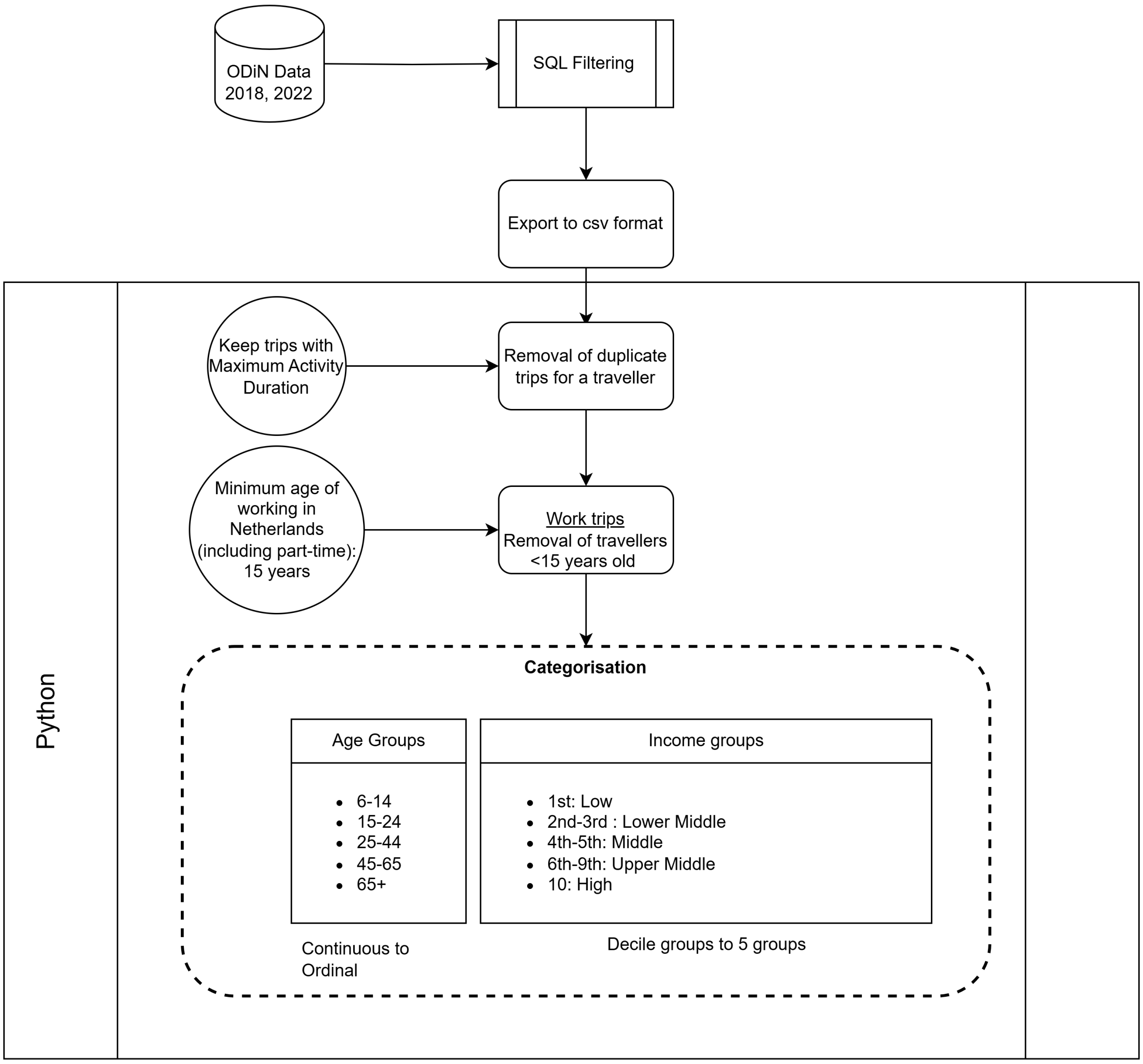


Figure 4.3: ODiN Data processing workflow

Figure 4.3 provides an overview of the steps involved in processing the 2018 and 2022 ODiN data. After using SQL queries to apply filters, as shown in Table 4.1, the filtered MRA region trip data is then processed further in Python by importing it as a CSV file using the following steps:

1. **Removing duplicated trips:** Generally, people have only one primary location for work and education, where they spend a significant part of their day. Hence, to determine the primary arrival destination, it is assumed that the primary destination is where the person has the maximum activity duration. Following this assumption, for each traveler, only the trip with the maximum duration is retained in the case of multiple trips for trip purposes.
2. **Minimum age limit (Work Trips):** In the Netherlands, including part-time work, the minimum age requirement is 15 years for 2018 and 2022. Consequently, travelers younger than 15 years are excluded from the dataset, in line with the Netherlands minimum working age (Business.gov.nl, 2024). This step ensures that only relevant trips are considered, adhering to the legal working age for the years in focus of this research.
3. **Categorisation of Variables** The age data for travelers recorded in the survey is a continuous variable. Interaction effects for age are included by creating dummy variables for different age, which will then be used to include interaction with the travel impedance (travel time). This approach avoids assuming a linear or monotonic relationship, and simplifies forecasting by allowing predictions based on categories rather than

continuous age values (Bhat et al., 1998). To convert the continuous age variable into categorical variables, appropriate groupings for age are needed to prevent the loss of information. Five categories are defined for age groups using the same ranges used by the Centraal Bureau voor de Statistiek (CBS) in the Netherlands. CBS reports key demographic, household, income, housing, and accessibility statistics by postal code and grid size (e.g., 100m x 100m, 500m x 500m) across the Netherlands (Centraal Bureau voor de Statistiek (CBS), 2023). Given that ODiN data includes only travelers aged 6 years and older, the following age groups are selected: (a) 6–14 years, (b) 15–25 years, (c) 25–44 years, (d) 45–64 years, and (e) 65 years and above.

For income groups, the ODiN data categorizes travelers by deciles of standard household disposable income. To make the model more parsimonious, these ten groups need to be regrouped into fewer groups while ensuring minimal loss of information owing to aggregation into fewer groups. Standardized household income, initially provided in deciles, is re-categorized into five broader groups (table 4.2).
Regrouping is guided by annual median standardized household disposable income levels for 2018 and 2022 in the Netherlands, as reported by the OECD Organisation for Economic Co-operation and Development (OECD) (2023), and further refined with CBS's reported median standardized household disposable income data for each decile group (Centraal Bureau voor de Statistiek (CBS), 2024). CBS labels the 1st decile as 'Low' and the 10th decile as 'High,' so these labels are preserved to minimize loss of information due to regrouping. Based on these considerations, Table 4.2 presents the restructured income groups aligned with OECD's reported standardized household disposable median income for the Netherlands for 2018 and 2022. The value in the bracket below the years indicates the standardized household disposable median income for the Netherlands for the respective years.

Table 4.2: Criteria for regrouping Decile Income groups

New Groups	**Decile Groups**	**% of Median Annual Standard Disposable Income Level**	**Income range per year** In Euros	
			2018 (27,000)	2022 (32,000)
Low	1st	≤50%	13,500	16,000
Lower Middle	2nd and 3rd	50% upto 90%	13,500 to 24,300	16,000 to 28,800
Middle	4th and 5th	90% upto 125%	24,300 to 33,750	28,800 to 40,000
Upper Middle	6th to 9th	125% to 250%	33,750 to 67,500	40,000 to 80,000
High	10th	≥250%	67,500	80,000

4.3.2. Employment data processing

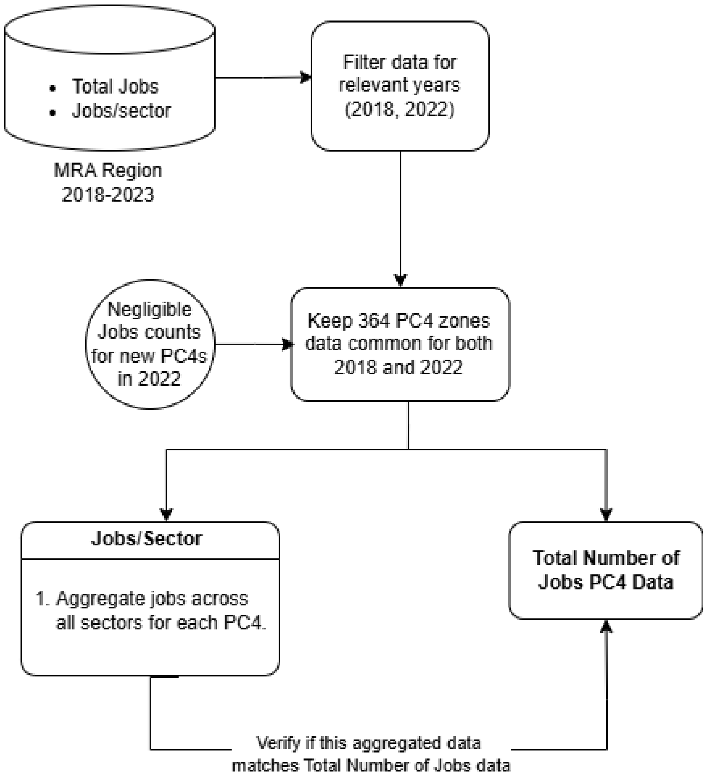


Figure 4.4: Employment Data processing workflow

The raw data, provided by O&S Amsterdam, includes all jobs and jobs per sector from 2018 to 2023. For this research, only data from 2018 and 2022 are within scope. This data is recorded annually on April 1st. For instance, the 2018 data corresponds to April 1, 2018. Given that the ODiN survey can occur at any time during the year, it is more accurate to choose the April 1, 2019, employment data to correspond with the ODiN survey data for 2018. A similar approach is applied for selecting data from 2023 to correspond with 2022 ODiN Survey data. Two separate dataframes are created; one for 2018 and another for 2022, each filtered for the relevant years.

As seen in Figure 4.4, the Jobs/sector data is summed up across all sectors and aggregated on a PC4 zone level and verified with the PC4 level Total number of Jobs data to ensure the data is processed correctly. Owing to GDPR regulations, very low numbers (<5) are converted to zero by the O&S team at Amsterdam. Therefore, the numbers would not exactly match but would be in close proximity to each other.
One issue identified with postal codes, where new codes are added to municipalities within the MRA region in 2023, leading to a discrepancy in the number of postal code zones (PC4) between 2018 and 2022. Specifically, there were 364 PC4 zones in 2018, which increased to 371 zones in 2022. Upon examining the data, it was noted that the number of jobs in these newly added PC4 zones in 2023 is very low. These zones exist only in the 2023 data and have negligible job numbers; hence, they are excluded. To ensure consistency, only the 364 PC4 zones common between 2018 and 2022 are considered.

4.3.3. Education data processing

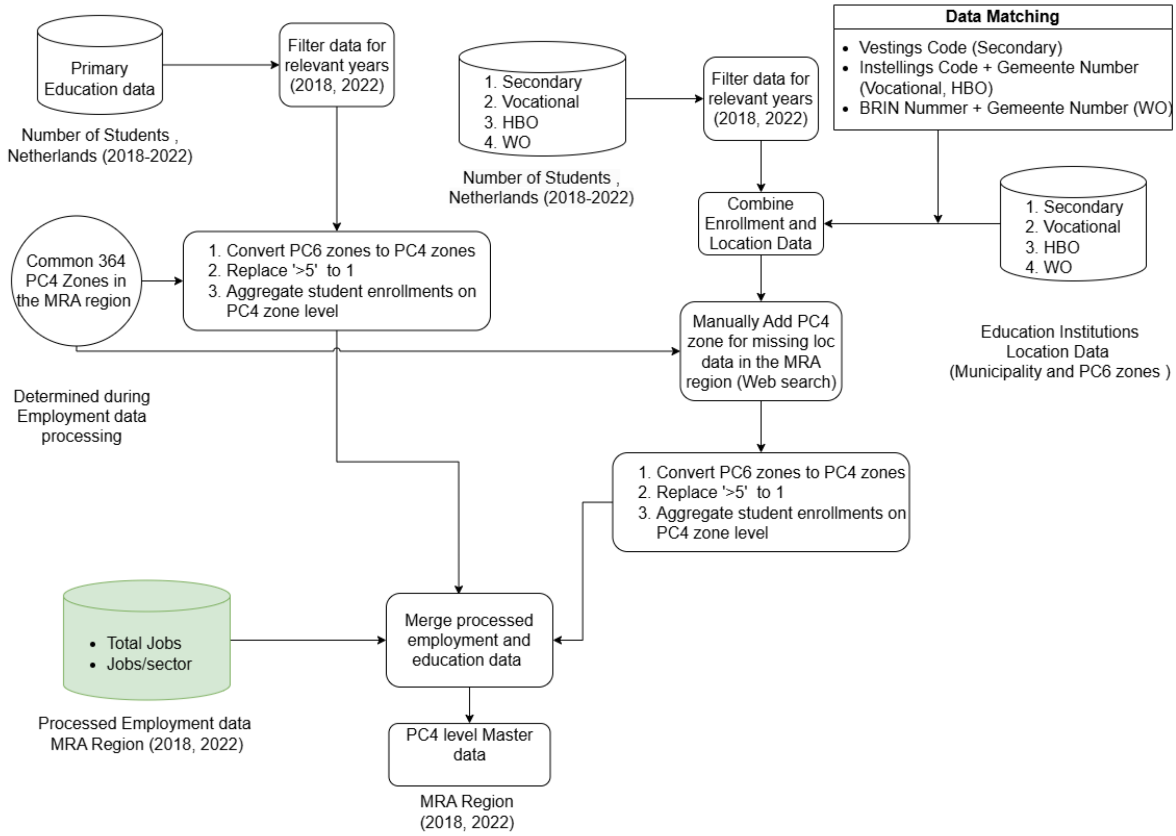


Figure 4.5: Education data Processing

Figure 4.5 provides an overview of the workflow of educational data processing. As observed, these are different datasets pertaining to different levels of education, as defined in the Dutch education system and shown in figure E.1 in Appendix E. They have different data structures and are processed differently according to the requirements for estimating destination choice models for education trip purpose. All student numbers for an academic year corresponded to the number of students in November. For example, the 2018 data would contain the total number of students at the institution as of November 2018.

Primary Education enrollments

For the relevant years (2018 and 2022), the data initially provided at the more granular PC6 postal code level is aggregated to the PC4 level. The number of students is grouped and summed up according to the PC4 postal codes.

Secondary enrollments

1. Data Source and Location Matching: Unlike primary education, secondary education enrollment data does not contain PC4 or PC6 location codes. To resolve this, an additional location file from the DUO website is used to determine the geographic location of secondary schools. The data is then matched using the Vestigings code (institution code) or Vestigings Nummer (institution number).
2. Handling Missing Location Data: This matching process may result in some missing (NaN) PC4 codes due to the absence of corresponding location data. To address this, institutions with NaN values in PC4 codes are filtered to retain only those located within the Metropolitan Region Amsterdam (MRA). For the remaining institutions without PC4 codes but located in the MRA region, the PC4 code is manually retrieved using Web Search based on the institution name and municipality information provided in the data.

The final step of aggregating the data at the PC4 level is similar to that of the primary education data.

Vocational enrollments

Vocational education data is processed in a manner similar to secondary education. Enrollment data is matched using the Instellings Code (institution code). Student counts of less than five are reported as "<5" due to the GDPR. In such cases, these values are set to "1".

HBO enrollments

The HBO education data is processed similarly to that of vocational education. However, because of the possibility of multiple branches of the same institution sharing the same Instellings Code, enrolment and location data are matched not only by the Instellings Code but also by the Gemeente (municipality) number and name to ensure that student enrollments are correctly assigned to the appropriate PC4 code.

WO (University Education):

WO education data is processed similarly to HBO education data, with the BRIN Nummer (a unique identifier for institutions) used as the matching key.

After processing and aggregating the educational data at the PC4 level for each educational category, they are merged with the employment data processed previously at the PC4 level. This creates a master dataset for 2018 and 2022, allowing for all job and education enrolment data for each of the 364 PC4 zones in the MRA region.

4.4. Availability of Destinations

Table 4.3: Number of available destination alternatives and observed trips after data processing

Trip Purpose	2018		2022	
	Available PC4 zone destinations	No. of Trips	Available PC4 zone destinations	No. of Trips
Home-Based Maintenance	350	1,874	355	1,851
Work	363	2,680	363	1,903
Secondary and above Education	125	364	129	276
Primary Education	278	693	282	395

Table 4.3 presents the number of available destination zones for each trip purpose out of the total 364 PC4 zones and the number of observed trips arriving in these available MRA PC4 zones for 2018 and 2022, after processing all four data sources.

In transport modeling, particularly in destination choice models, it is common practice to apply a feasibility criterion to reduce the size of the choice set. This can be achieved by considering the compatibility between trip types and land-use characteristics to ensure that only realistic and relevant alternatives are included in the model (Bernardin et al., 2018). Travelers are unlikely to consider destinations that do not offer opportunities relevant to their trip purpose. This prevents misrepresentation of destination attractiveness. In addition, choice set reduction reduces the computational resources required for model estimation. Hence, in this study, the number of destination PC4 zones available to travelers depends on the availability of opportunities for the trip type. This means excluding zones with no retail employment for HBM trips, zones with no employment as possible work locations, and zones with no enrollment at the relevant education level as possible locations for education trips. Consequently, the observed trips in the ODiN data is further filtered to include only arrivals in these available PC4 zones.

4.5. Implications of Modelling Assumptions during Data processing

The assumptions made in the modeling process have significant implications for estimating destination choice models for work, education, and home-based maintenance (HBM) trips. First, converting continuous variables, such as age and income, into categorical variables simplifies the models and avoids assuming linear relationships; however, this can obscure subtle variations within categories. This aggregation may lead to a loss of important behavioral insights, as travelers within the same age or income category might have different preferences or constraints that affect their destination choices. Second, the assumption for intrazonal travel times, assuming them to be half the travel time to the nearest neighboring zone for each destination zone, is a rough approximation that may not accurately reflect actual travel times within zones. This can lead to inaccuracies in modeling travel impedance, particularly for HBM activities where intrazonal trips are dominant.

Finally, for work and education trips, it is essential to determine the traveler's actual workplace or educational institution. However, the ODiN travel survey data does not explicitly specify whether the recorded destination zone is the traveler's place of work or education. To address this issue, certain assumptions must be made. Generally, individuals have only one primary location for work or education, where they spend a significant portion of their day. Therefore, in this study, it was assumed that the destination with the maximum activity duration represents the primary work or education location. Accordingly, for each traveler with multiple trips to different destinations for the same trip purpose, only the trip with the longest duration was retained. However, it is possible that this location might not actually be the workplace or the primary place of education. Travelers may spend significant time at locations that are not their workplace or educational institution such as client sites, training centers, conferences, cafes or other long-duration activities unrelated to their primary work or education location and record them as work or education trip purpose. As a result, this could lead to misclassification of destinations, potentially affecting the accuracy of destination choice models. The model might incorrectly estimate the attractiveness of certain zones or fail to capture the true patterns of work and education trips [9]

4.6. Utility Specification

This section presents the utility specification formulated for the three trip purpose: HBM, Work, and education (Primary and secondary & above education level choices)

Table 4.4 summarizes the motivation for including the variables in each specification based on the findings from the literature.

Additionally, for each trip purpose, the zonal size parameter was estimated here. In transport modeling, it is common practice to fix the zonal size parameter to 1. However, in this study, the parameter is estimated as it is used later in determining the choice set size for sampling alternatives. This is because the approach is based on the stability of the parameter, as described later in Section 4.7.1

[9]While the assumption of retaining only the trip with the maximum activity duration is appropriate for work and education trips, as individuals generally have one primary location for these purposes, it does not hold true for home-based maintenance trips, where multiple destinations may be visited within a single day. Initially, due to oversight, I filtered the data under the assumption that only one observation per traveler per trip purpose was needed for HBM trips as well, overlooking Pandas Biogeme's capability to handle multiple observations per traveler using PanelObs = True. Unfortunately, I identified this mistake late in the process, and because of time constraints, it was not possible to perform all validation steps with the correct data structure at the time of submission of the final version of this document. However, I conducted a quick informal parameter estimation, which showed that the parameter estimates and t-test values for most significant parameters, especially SC&AE, and model performance in both years, did not differ significantly from the results presented in the Results chapter. Therefore, the general storyline remains the same.

Table 4.4: Motivation of including interaction effects of various traveler characteristics for each trip purpose.

Trip Type	Interaction effects with travel time included	Motivation
Work	Gender, Age Group and income groups	Bhat et al. (1998) explored sociodemographic interactions with travel impedance for HBM and Home based Work trips in the Boston Metropolitan Area, USA, finding that older adults and women were more sensitive to travel impedance for work trips, while higher-income travelers were more willing to travel longer distances for work.
HBM		
Education Trips	Migration Background, Income Groups	Previous studies such as de Boer and Blijie (2006) in Zwijndrecht, Netherlands, found that for primary school choice, sensitivity to travel time for school trips varies by income and migration background. Higher-income families are sensitive to travel barriers but willing to travel further for specific school attributes, while lower-income families are less affected by distance. Non-Western families are sensitive to travel barriers but less focused on school type or quality. In secondary education, affluent Native Dutch students prefer nearby schools, while those in lower socioeconomic areas travel farther. Non-Western students prefer local schools with peers from similar backgrounds (van Welie et al., 2013) . In the present model, income groups are used in place of socioeconomic status, and only the interaction effect for the migration background of students is considered, as there is no available data on the distribution of migration backgrounds within the schools in PC4 zones.

Home-Based Maintenance Trips

T_{ij} represents the utility of traveler living in PC4 zone i from choosing PC4 j to perform maintenance trips. It is as follows:

$$T_{ij} = \beta_m \ln(M_j) + \beta_c \ln(c_{ij}) + \sum_{n=1}^{4} \beta_n \cdot A_n \cdot \ln(c_{ij}) + \beta_f G_f \cdot \ln(c_{ij}) + \sum_{i=1}^{4} \beta_i \cdot I_i \cdot \ln(c_{ij}) + \beta_{A_j} \ln(\sum_{k \neq j} \frac{M_k}{c_{jk}}) \quad (4.1)$$

Here, M_j is the number of relevant employments in PC4 zone j. The sectors from which job counts are considered relevant are Wholesale & Retail, Financial Institutions, Utilities, Government, Health & welfare, and Other services (6 sectors in total). c_{ij} is the Travel cost between zones i and j with cost sensitivity parameter β_c

A_n is the age group n interacting with travel cost c_{ij} with β_n capturing the interaction effect. Four age groups starting from 6-14 are included, with the 65+ age group as the reference level for dummy coding. Next, for gender interactions with travel time, G_f is the Gender Binary variable. A value of 1 indicates that the traveler is female. β_f is the female gender parameter, with male as the reference level.
For the income interaction with travel time, I_i is the income group i interacting with travel cost c_{ij}. β_i captures the interaction effect. The income groups are based on the five income groups presented in Table 4.2 with the middle-income group as the reference level for dummy coding.

Lastly, for including SC&AE, c_{jk} is the travel time between zones j and k with spatial structural parameters β_{A_j} accounting for SC&AE for destination zone j.

Work Location choice
W_{ij} represents the utility of travelers living in PC4 zone i who choose to work in PC4 j. It is as follows:

$$W_{ij} = \beta_j \ln(J_j) + \beta_c \ln(c_{ij}) + \sum_{n=1}^{3} \beta_n \cdot A_n \cdot \ln(c_{ij}) + \beta_f G_f \cdot \ln(c_{ij}) + \sum_{i=1}^{4} \beta_i \cdot I_i \cdot \ln(c_{ij}) + \beta_{A_j} \ln(\sum_{k \neq j} \frac{J_k}{c_{jk}}) \quad (4.2)$$

In this utility specification, J_j is the total number of jobs in PC4 zone j. For age groups, with 15 years being the minimum age for working (including part-time) in the Netherlands (Business.gov.nl, 2024), the age group 6-14 is excluded.
Apart from these differences, the remaining variables used to explain destination choices for work trips are the same as for HBM trips (Eq. 4.1).

Education location choice
None of the studies reviewed in Section 2.3 considered age as an explanatory variable to explain destination choices for education trips. Moreover, Primary education is compulsory in The Netherlands (Nuffic, 2024). Hence, in this thesis, age is not considered to play a significant role in explaining destination choices for primary education. For secondary and above education level trips, in this study, age is not considered to play a significant role, as it is more likely that the education level available and the courses offered at the institutions attract trips to the zones.

- Primary School location choice

  For this trip purpose, the ODiN survey data is filtered to consist of travelers below 15 years of age Hence, they have *Opleiding* (Attained Education Level) as ‘Not Requested’ as per the survey protocol for ODiN. Therefore, it is assumed that these travelers with trip ’Motive’ as Education and ’Purpose’ as Educational course in ODiN data are performing trips for Primary education.

  S^p_{ij} represents the utility of traveler living in PC4 zone i from choosing PC4 j to perform primary education level trips. It is as follows:

$$S^p_{ij} = \beta_p \ln(P_j) + \beta_c \ln(c_{ij}) + \sum_{m=1}^{2} \beta_m \cdot M_m \cdot \ln(c_{ij}) + \sum_{i=1}^{4} \beta_i \cdot I_i \cdot \ln(c_{ij}) + \beta_{A_j} \ln(\sum_{k \neq j} \frac{P_k}{c_{jk}}) \quad (4.3)$$

  Here, P_j is the number of primary education enrollments in PC4 zone j. M_m is the Migration background m interacting with travel cost c_{ij}. β_m captures the interaction effects. Travelers are divided into three groups: Dutch (reference level), Western, and Non-Western Migration. These three groups are dummy-coded into two levels.
  The rest of the variables are similar to the ones in HBM trips (eq. 4.1)

- Secondary and Above Education Level location choice

  It consists of travelers who completed ‘Basic education’ (primary) and above in the *Opleiding* column of the ODiN data column.

  For education trips, the number of trips available for each education level after the primary level was very low compared to the rest of the trips (table 4.3). Hence, these trips were combined to form the purpose of ’Secondary education and above.’ In these trips, travelers with attained education levels (*Opleiding*) in ODiN data at secondary (VMBO, HAVO, VWO), vocational (MBO), and higher education (HBO, WO).

  S^s_{ij} represents the utility of traveler living in PC4 zone i from choosing PC4 j to perform secondary education level trips. It is as follows:

$$S^s_{ij} = \beta_s \ln(S_j) + \beta_c \ln(c_{ij}) + \sum_{m=1}^{2} \beta_m \cdot M_m \cdot \ln(c_{ij}) + \sum_{i=1}^{4} \beta_i \cdot I_i \cdot \ln(c_{ij}) + \beta_{A_j} \ln(\sum_{k \neq j} \frac{S_k}{c_{jk}}) \quad (4.4)$$

Here, the utility specification is the same as that for primary education trips, except for S_j as the zone size measure to ensure that only enrollments at the relevant education level are considered for this trip purpose. S_j is the number of secondary and higher education enrollments in PC4 zone j. These include Secondary Education (VMBO, HAVO, and VWO), vocational education (MBO), and higher education (HBO and WO). Figure E.1 in Appendix E provides an overview of how these levels are related.

4.7. Sampling of Destination alternatives

Including all zones in the study area in the choice set is neither a realistic representation of how travelers choose destinations nor practical, in terms of computational efficiency. It is widely understood that travelers do not evaluate such a large number of alternatives when selecting a destination. They tend to automatically eliminate destinations that are too far from their origin zone (Bernardin et al., 2018). Hence, the sampling of destinations is required. Sampling alternatives requires making two key decisions. First, determining the appropriate choice set size, and second, selecting an adequate sampling method.

4.7.1. Determination of Choice set size

Regarding the appropriate choice set size, Guevara et al. (2016) presented a method based on a Monte Carlo experiment to determine the set size based on the stability of parameter estimates, both the average and standard deviation, suitable for various models, including the RRM, MEV logit mixture, and logit models. Specifically, Guevara et al. (2016) varied the choice set sizes ($\tilde{J}$) and for each $\tilde{J}$, sampled K times. They then estimated the mean parameter values ($\bar{\hat{\beta}}$) and their standard deviations as follows:

$$\bar{\hat{\beta}} = \frac{1}{K}\sum_{k=1}^{K}\hat{\beta}_k(\tilde{J}) \quad \text{and} \quad \hat{\sigma}_{\hat{\beta}} = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}\left(\hat{\beta}_k - \bar{\hat{\beta}}\right)^2}$$

This approach helps identify an appropriate choice set size for the given travel survey data based on the stability of the parameter estimates. This was also applied by Mauad and Isler (2024) in a destination choice model for home-based work trips in São Paulo, Brazil. Both Guevara et al. (2016) and Mauad and Isler (2024) use 30 iterations (K = 30) for each sample size. This research too adopts the same methodology.
In this study, the choice set size is varied from 5 to 50 destination alternatives. For each choice set size, the mean beta values and standard deviations are calculated across 30 iterations for the model specifications enlisted in Section 4.6. Although both statistically significant and non-significant parameters are included in the 30 iterations (using different random seeds), only the statistically significant parameters from the full choice set are plotted in the results analysis (Chapter 5).

4.7.2. Stratified Importance Sampling

Concerning the second decision on the sampling method, this study uses a variant of the Stratified Importance Sampling (SIS) method adapted from Bradley et al. (1998). As illustrated in figure 4.6, for each origin zone, destinations are chosen based on their distance from the origin PC4 zone and the destination's attraction size relevant to the trip purpose.

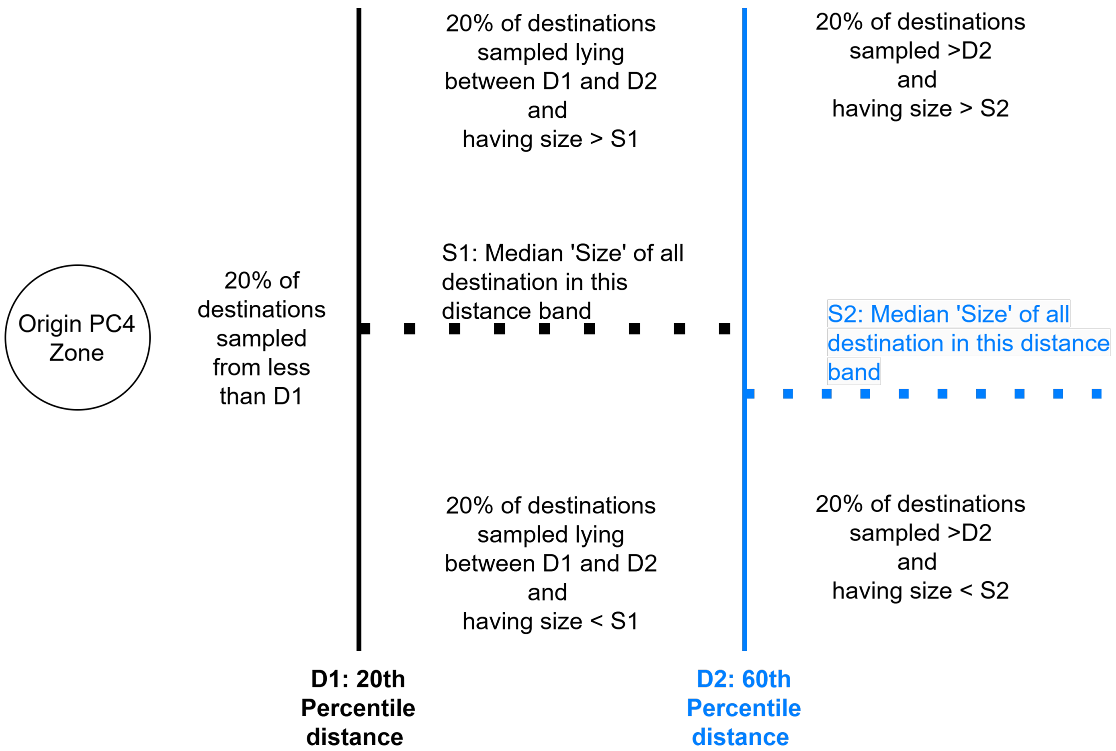


Figure 4.6: Stratified Importance sampling for a given origin PC4 zone representing the actual distribution of trip destinations based on both distance and 'size' of destinations

In this stratified sampling method, adapted from Bradley et al. (1998), destinations are chosen based on their distance from an origin PC4 zone and their attraction size, such as employment or enrollment, depending on the trip purpose. The sampling approach is divided into distinct distance thresholds: the 20th percentile distance, labeled D1, and the 60th percentile distance, labeled D2. D1 represents closer destinations, whereas D2 represents an intermediate range. Zones beyond D2 are considered to be the farthest destinations.
Attraction size further stratifies the destinations within these distance ranges. Median employment levels, referred to as S1 and S2, are calculated separately for destination zones between D1 and D2, and for those beyond D2. Attraction size categories vary by trip purpose, using total employment for work, relevant education level enrollments for education purposes, and retail & service employment for HBM trips. S1 represents the median employment size within the D1 to D2 distance band, whereas S2 is the median size for zones farther than D2.
For each sample size, samples are drawn from different strata to mirror the actual distribution of tour destinations in terms of distance and size. Specifically, 20% of the destinations are chosen from zones closer than D1, 20% from zones between D1 and D2 with employment less than S1, 20% from the same distance band with employment greater than S1, 20% from zones beyond D2 with employment less than S2, and 20% from the farthest zones with employment exceeding S2. By reflecting proximity and attractiveness, this variant of SIS ensures one possible realistic representation of destination choice sampling.

4.8. Assessing Impact of SC&AE on temporal transferability

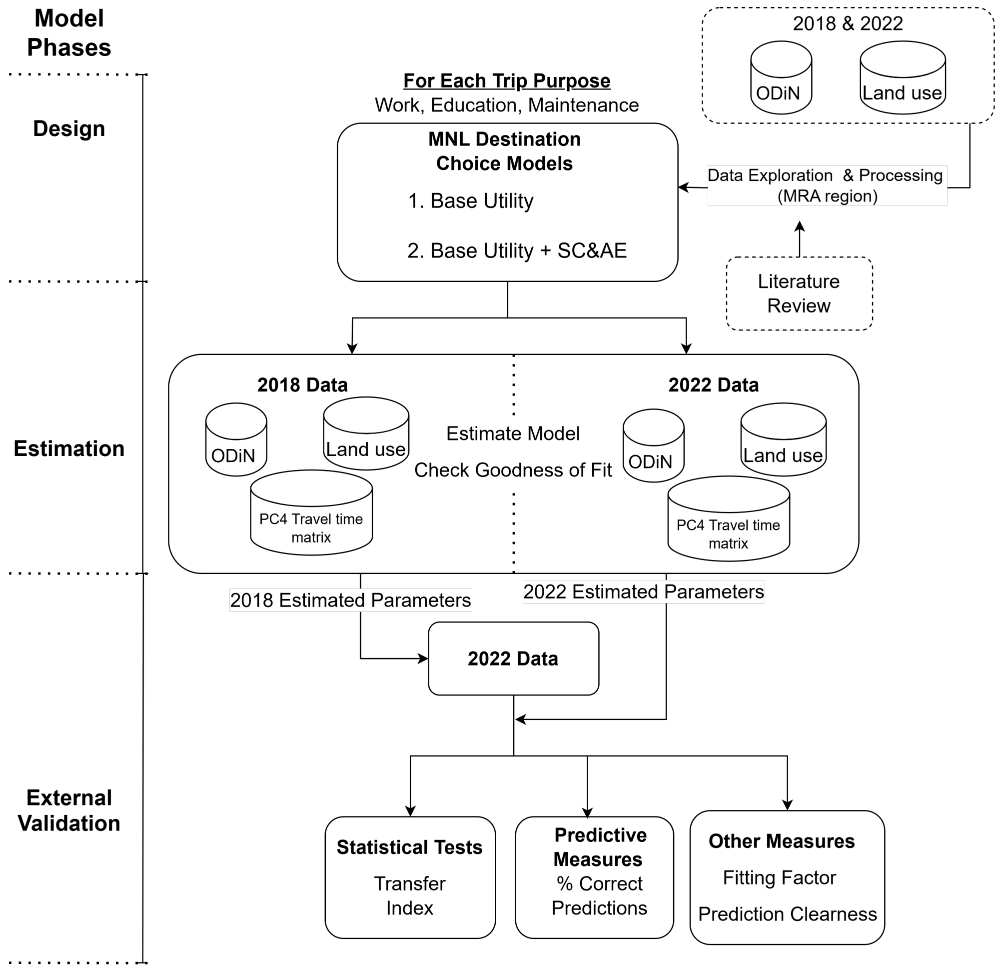


Figure 4.7: Methodology Overview of Validation Phase (SQ4)

Figure 4.7 illustrates the core methodology used to assess the impact of SC&AE on the temporal transferability of destination choice models. The base utility specification designed in section 4.6 by using all the variables selected in Section 4.2 , except for the SC&AE parameter. Two models (one with base utility and the other inclusive of SC&AE) are estimated using data from 2018 (base year) and 2022 (forecast year) for each trip purpose. The 2018 model parameters are then applied to 2022 data and compared with models re-estimated on 2022 data. The comparison involves four measures across three categories: (1) Transfer Index, (2) Percentage of correct predictions and other measures, (3) Discriminative ability, and (4) Fitting Factor.

4.8.1. Selection of Performance Indicators

The parameter estimation of discrete choice models is based on the Maximum likelihood principle, which is a statistical method to estimate model parameters by finding values that maximize the likelihood of observing the given data (Bunch, 1987). Hence, log-likelihoods are commonly used to assess the explanatory power of discrete choice models. Therefore, the first transferability test is a statistical test (Transfer Index), whose calculation is also based on the log-likelihoods of the models estimated across the forecast horizon.

Because these models are used as predictive tools, it is important to quantify or translate how statistical test performance translates into performance in terms of the model's practical use. Therefore, the second category to assess the impact is Predictive measures such as % Correct predictions.

However, discrete choice models are probabilistic models, not deterministic. Discrete choice models, such as the Multinomial Logit, predict the probability of choosing an alternative, not the actual choice itself (Hauser, 1978). Hence, other measures, such as the Fitting factor and Prediction clarity, are required to assess the quality of the probabilistic predictions.

Transfer Index (TI)

Statistical tests in the early literature were used to test the extent of the model transferability (Fox et al., 2014). This includes the Transfer Index (TI). Developed by Koppelman and Wilmot (1982), this index measures the predictive accuracy of a transferred model compared to a locally estimated model. In the context of this research, the 2018 ODiN Data is the Base year sample, and the 2022 ODiN Data is the Transfer year sample and disaggregate equivalent of a gravity model is the simple reference model. In this study, the disaggregate equivalent of a gravity model was chosen as a reference model because destination choice models are often considered a more realistic alternative to traditional aggregate approaches such as the gravity model because they are based on theories of individual behavior and do not rely on physical analogies and MNL is the simplest form of the destination choice model (de Dios Ortúzar and Willumsen, 2011). The utility specification of the disaggregate equivalent of a gravity model consists of only the travel impedance and zonal size measure. Hereafter, 2018 is used to reference the base year sample and 2022 for the transfer year sample.

A transferred model, estimated using the 2018 sample data, is then applied to the 2022 sample data. The index quantifies the extent to which the transferred model (using parameters estimated from the 2018 sample) outperforms the locally estimated simple reference model (using parameters estimated from the 2022 sample) for the choices observed in 2022. The upper limit of TI is one, indicating that the accuracy of the transferred model matches that of the 2022 locally estimated model. If the index shows negative values, the transferred model performs worse than the 2022 estimated simple reference model (Parady et al., 2021). Although this index does not strictly determine a pass/fail outcome, it offers a comparative measure of the model's transferability (Fox et al., 2014). The index is calculated as follows:

$$TI_{22}(M_{18}) = \frac{LL_{22}(M_{18}) - LL_{22}(M_{22}^{ref})}{LL_{22}(M_{22}) - LL_{22}(M_{22}^{ref})} \tag{4.5}$$

where:
$LL_{22}(M_{22}^{ref})$: Log-likelihood of the simplistic reference model for the 2022, in this case, A disaggregate equivalent of a gravity model in this study

$LL_{22}(M_{18})$: Log-likelihood for the model (using parameters estimated on the 2018 ODiN sample) estimated on the 2022 ODiN data

$LL_{22}(M_{22})$: Log-likelihood for the model estimated directly on the 2022 ODiN data.

% Correct Predictions

This measure evaluates the accuracy of the model by calculating the ratio of correct predictions to the total number of observations, expressed as a percentage. The alternative to which the model assigns the highest probability among all alternatives in the choice set is the predicted choice (Parady et al., 2021).

$$\frac{\text{Number of correct predictions}}{\text{Total number of observations}} \times 100 \tag{4.6}$$

$$\frac{\sum_{n_t=1}^{N_t} \hat{y} = y}{N_t} \times 100 \tag{4.7}$$

where:
y is the observed choice in 2022 ODiN data (validation sample)
$\hat{y}$ is the predicted choice. i.e., the choice to which the model assigns the highest probability in the choice set
N_t is the total number of observations

Prediction Clearness

One of the key limitations of the above % Correct Predictions is its inability to account for the model's discriminative ability in its evaluation (De Luca and Cantarella, 2009). For example, in a four-choice set scenario, where the second choice is the observed choice in the data, two models giving fractions of (15%, 35%, 25%, 25%) or (5% 85%, 5%, 5%) are considered the same on this indicator. However, it is clear that the 2nd model is much better at distinguishing the right choice. Therefore, this study includes this indicator to overcome this limitation. Considering a threshold for assigned probability, this indicator has three measures based on binary logic: (a) % Clearly right (%CR), where the observed choice is assigned a probability above the threshold. (b) % Clearly wrong (%CW), where any other alternative in the choice set, other than the observed choice, is assigned a probability above the threshold. (c) % Unclear (%UC), counting the choices assigned neither clearly right nor clearly wrong (De Luca and Cantarella, 2009). These three indicators are given as follows:

$$\%CR = \frac{100}{N_t} \sum_{n_t=1}^{N_t} CR_{n_t} \tag{4.8}$$

where:

$$CR_{n_t} = \begin{cases} 1 & \text{if } \hat{P}(y_{observed}) > t \\ 0 & \text{otherwise} \end{cases} \tag{4.9}$$

$$\%CR = \frac{100}{N_t} \sum_{n_t=1}^{N_t} CW_{n_t} \tag{4.10}$$

where:

$$CW_{n_t} = \begin{cases} 1 & \text{if } \hat{P}(!y_{observed}) > t \\ 0 & \text{otherwise} \end{cases} \tag{4.11}$$

$$\%UC = 100 - (\%CR + \%CW) \tag{4.12}$$

Setting the threshold t for the clearness of prediction is a modelling decision. In research, the general thumb of the rule is to keep the threshold "considerably" larger than c^{-1}, c being the choice set size. However, there is no clear consensus on a threshold that is "considerably larger" (Parady et al., 2021). It seems logical that the idea behind setting a threshold sufficiently high for a model is that not more than one choice should be assigned a probability above the threshold. A general recommendation by Parady et al. (2021), in case there is no clear consensus on a definite threshold, is to plot the results over a range of threshold values, as done by De Luca and Cantarella (2009). In this study, a threshold range of 40% to 90% is used.

Fitting Factor (FF)

Considering the sample size, FF measures the probability that a model assigns to the observed choice on average. It has an upper bound of 1, indicating that, on average, the model assigns a probability of 1 to the observed choice, hence perfectly forecasting all choices in the sample (De Luca and Cantarella, 2009). If $\hat{P}(y_{observed})$ is the assigned estimated probability of the observed choice y for observation n_t then:

$$FF = \frac{100}{N_t} \sum_{n_t=1}^{N_t} \hat{P}(y_{observed}) \tag{4.13}$$

4.8.2. Comparision of Impact Across Full and Sampled Choice set (Random and Stratified Importance Sampling)

As noted in Section 4.7, sampling in destination choice modelling is essential for a realistic behavioral representation and computational feasibility. Consequently, sampling errors are inevitable. To thoroughly understand the impact of SC&AE), this study also applies the framework illustrated in Figure 4.7 not only to models estimated using SIS, but also to models estimated with the same choice set size using random sampling. Random sampling is one of the most commonly used methods in destination choice modelling, in addition to various variants of SIS (Kim and Lee, 2017). This approach provides insight into how the impact of SC&AE on temporal transferability varies with the sampling method. Moreover, because random sampling is known to be an unrealistic method (assigning equal probability to all possible destination alternatives in the study area), it allows us to explore how selecting an incorrect sampling method affects the temporal transferability of DCMs, a factor not previously explored in research.

In addition to sampled models, a full-choice estimation is performed for the three trip purposes, considering all available alternatives without any sampling. This is done to establish an unbiased benchmark, free from sampling-induced errors, against which the performance of sampled models could be compared. By eliminating sampling biases, the full-choice estimation offers a clearer view of the true effects of SC&AE. It provides a baseline for assessing how different sampling methods influence parameter estimates and temporal transferability.

5
Results

This chapter presents the results of the impact of SC&AE on the temporal transferability of destination choice models across different trip purposes, including home-based maintenance (HBM), work locations, and education locations. The analysis uses multiple performance indicators, such as the Transfer Index, Fitting Factor, Percent Correct Predictions, and Clarity Analysis graphs, to evaluate the impact. Each trip purpose has a separate section, each with two subsections each. These include (a) Parameter estimates and (b) Results for the four indicators to evaluate the impact of SC&AE on temporal transferability. Comparisons are made between Full and sampled choice sets (Random and stratified Importance sampling).

Additionally, at the start of each section, before elaborating on these two aspects for each trip purpose, the variation in the average beta values and standard deviation to determine the choice set size is illustrated. The below table 5.1 summarizes the chosen choice set size for each trip purpose from the available destinations

Table 5.1: Number of available destination alternative and choice set size across trip purpose

Trip Purpose	Number of available PC4 zone destinations		Choice set size
	2018	2022	
HBM	350	355	40
Work	363	363	45
Secondary and above Education	125	129	45
Primary Education	278	282	278

Table 5.1 presents the number of available PC4 destination zones with nonzero zonal size measures and the corresponding choice set chosen for each trip purpose. Initially, for trips, the model was estimated on the full choice set (including all available destinations), before estimating the parameters for the corresponding choice set size determined for sampling. As the full choice set would include all destinations, this was done to check whether the parameter estimate was statistically significant for each trip purpose. Moreover, the SC&AE parameter remained relatively the same, with similar statistical significance across the full and sampled choice sets. Hence, this approach helped save time by avoiding determining the choice set size for sampling when the SC&AE parameter was found to be statistically insignificant for the full choice set.
For primary education, the SC&AE parameter was found to be statistically insignificant in the 2018 data for the full choice set. Hence, the method of determining the choice set size was not used for primary education trips.

In addition, the results reveal several noteworthy findings, which are elaborated in section 5.5. This section provides an analysis on why the Transfer Index metric exaggerates the impact of SC&AE (Section 5.5.1), how accounting for SC&AE helps mitigate the loss in the Transfer Index value caused by parameter estimation errors introduced through sampling (Section 5.5.3) and the reasons behind the inconsistent performance of SC&AE for secondary and higher education level trip purpose (Section 5.5.2)

Finally, Section 5.6 concludes by discussing the implications of these results and their implications for destination choice models as predictive tools for policymaking.

5.1. Home Based Maintenance

The choice set size is determined based on the stability of the parameter estimates, both the average and the standard deviation, as described in Section 4.7.1. Figure 5.1 illustrates the variation in average beta values for Stratified Importance Sampling (SIS).

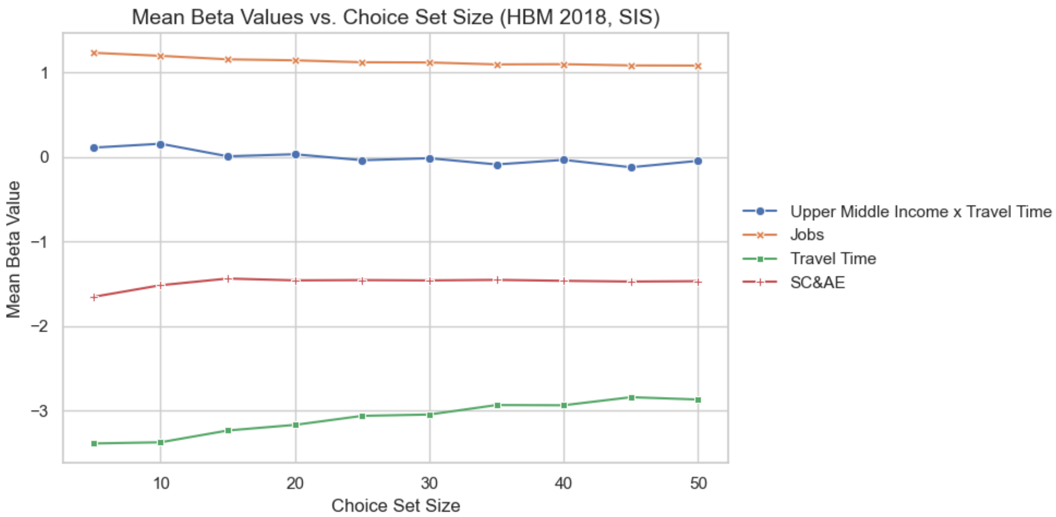


Figure 5.1: Mean Beta variation for Base Utility+SCAE (30 iterations/ choice set, HBM 2018, Stratified Importance Sampling)

As shown in Figures 5.1, Jobs, SC&AE, and the interaction between the upper-middle income group and travel time are stable across all choice set sizes. The Travel Time parameter stabilizes after a choice set size of 20, as indicated by the decreasing slope. By the choice set of size 30, SIS starts achieving values closer to -2.3, the parameter estimate for the full choice set.

Another important criterion for determining the appropriate choice set size is the stability of parameter standard deviations. By examining the standard deviations across different choice set sizes, we can identify the point at which stability is achieved. Because not all parameters will behave the same, Guevara et al. (2016) recommend prioritizing the stability of the least stable (or "worst behaving") parameters when determining the optimal choice set size. Figure 5.2 illustrates the standard deviations of each parameter across 30 iterations for various choice set sizes in the Base Utility+SC&AE model. By focusing on the parameters with the highest variability, the most appropriate choice set size can be identified based on when these parameters stabilize.

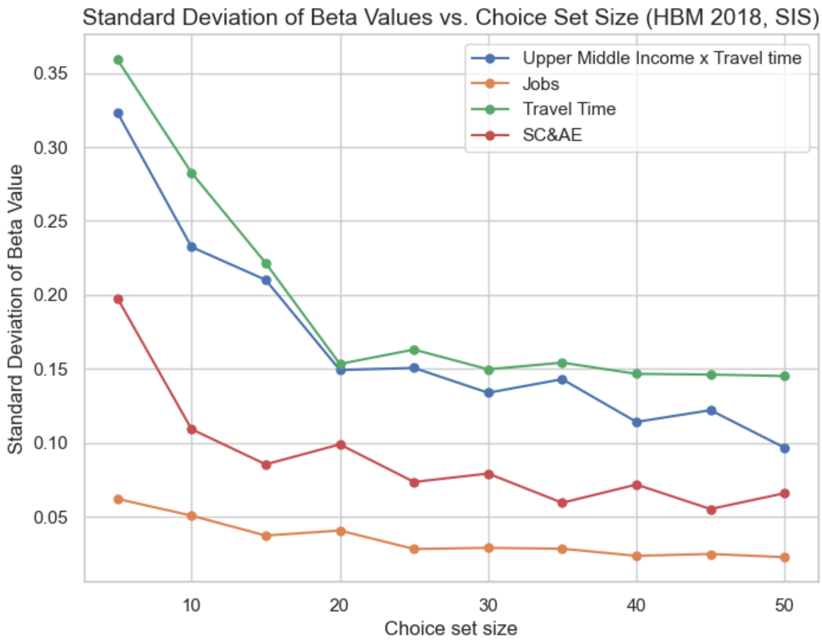


Figure 5.2: Standard Deviation variation for Base Utility + SCAE specification (30 iterations/ choice set, HBM 2018, Stratified Importance Sampling)

As shown in Figure 5.2, there is a significant reduction in the standard deviations of the Travel Time, the Income groups, and SC&AE parameters after the choice set size reaches 20, indicating a greater stability beyond this point. This trend aligns with the observations from the earlier plot of the average parameter values (Figure 5.1). Additionally, the standard deviations for the least stable parameter, travel time, decreased further after the choice set size reached 30. For the second least stable parameter, the income group interaction with travel time drops further for 40 alternatives. Considering the behavior of the second least stable parameter, the interaction of the Upper Middle-income group with travel time, a choice set size of 40 alternatives seems most appropriate for home-based maintenance trips

5.1.1. Parameter Estimates

Table 5.2 presents the parameter estimates for the full and sampled choice sets.

Table 5.2: Estimated Betas, HBM (2018, 2022 ODiN), Full vs. 40 Sampled Choice sets (RS and SIS)

Estimated Betas, HBM (2018, 2022 ODiN), Full vs 40 Alt Choice set (RS and SIS)							
		2018			**2022**		
		Full	**40 RS**	**40 SIS**	**Full**	**40 RS**	**40 SIS**
Note	Name	Parameter Values (t-test values)			Parameter Values (t-test values)		
Parameters for logarithmic values	Jobs	0.98 (26.42)	1.08 (22.92)	1.11 (23.24)	0.94 (24.91)	1.02 (19.73)	1.01 (21.75)
	SCAE	-1.40 (-10.46)	-1.35 (-7.17)	-1.60 (-8.56)	-1.69 (-13.70)	-1.66 (-9.19)	-1.77 (-10.55)
	Travel time	-2.30 (-34.37)	-3.10 (-17.36)	-2.74 (-16.61)	-2.41 (-38.32)	-2.95 (-18.39)	-2.75 (-18.85)
	Income Groups						
Middle Income Group as reference level	Lower Middle Income x Travel time	-0.138 (-1.757)	0.052 (0.22)	-0.33 (-1.38)	-0.137 (-1.72)	-0.73 (-3.05)	-0.91 (-3.72)
	Upper Middle Income x Travel time	-0.197 (-2.99)	0.088 (0.486)	-0.120 (-0.76)	-0.153 (-2.39)	-0.566 (-3.29)	-0.53 (-3.45)
	rho square (null)	0.508	0.698	0.677	0.512	0.711	0.685

As shown in Table 5.2, all the estimates have the expected signs. Only the parameters for Jobs, Travel Time, upper-middle income group, and SC&AE are statistically significant at the 5% level.

For SC&AE, the negative parameter value suggests that spatial competition is the dominant factor, which is consistent with the findings of Bhat et al. (1998) for home-based shopping trips and Bernardin et al. (2009), who used single parameters to capture the net effect of SC&AE on destination attractiveness in similar contexts. Among the income groups, only the interaction between the upper-middle income group and travel time is significant, yielding a negative parameter. This suggests that, relative to the middle-income group, the upper-middle-income group exhibits a higher sensitivity to travel time for maintenance trips from home.

Comparing the parameter estimate values across the full and sampled choice sets, the SC&AE parameter value is relatively stable across the full and two sampling methods. It is likely due to its intrinsic structure, which includes information on Retail & service employment and Travel Time across all alternatives (equation 4.1), which does not vary much across full and different sampling destination methods to form the choice set. Considering the full choice set parameter value as the ground truth, as it contains all possible destinations and hence has no sampling bias, compared to RS, the SC&AE values for SIS seem to deviate more. It has a % change of 14% in 2018 compared to RS's 3.57%, and 4.73% in 2022 compared to RS's 1.77%.

Comparing the rest of the parameters, in contrast to the stability of SC&AE, the Travel Time parameter experiences a substantial increase. However, when compared to RS, SIS yields values for the Travel Time parameter that are closer to the full choice set in both years, with an increase of only approximately 19% (versus RS's 34%) in 2018, and 14% (versus RS's 22%) in 2022. The interaction parameter for the Upper-Middle Income groups with Travel Time also exhibits a significant change, increasing from -0.153 to -0.566, an increase of approximately 270%. Although SIS presents a slightly lower increase, it remains high at approximately 246%.

Regarding discrepancies in parameter estimates between the full and sampled choice sets, in 2018, one interaction term between the upper-middle-income group and Travel Time, which was statistically significant at the 5% confidence level (t-test > 1.96), becomes insignificant for both sampling methods. Conversely, in 2022, the interaction term for the lower-income group becomes statistically significant in both sampling methods despite being insignificant in the full choice set model. These discrepancies in terms of statistical significance suggest possible errors in parameter estimates due to sampling

Moreover, using sampling, the model's rho-squared value increased by 34-39% for both years, likely reflecting enhanced discriminatory power when limited to 40 alternatives.

5.1.2. Impact of SC&AE on temporal transferability

Transfer Index, Fitting Factor, and % Correct Predictions

Table 5.3: HBM trips: Performance Comparision on various indicators Full vs Sampled choice sets (RS and SIS)

HBM 2022 (using 2018 estimated parameters)				
		Full Choice Set		
Indicators	Base Utility	Base Utility + SCAE	% Improved	Absolute gain in TI value
TI	-0.44	0.82	286.36%	1.26
Fitting Factor	0.2016	0.2068	2.59%	-
% Correct Prediction	36.09%	36.14%	0.14%	-
		40 Alt Randomly Sampled		
TI	-1.21	0.32	126.45%	1.53
Fitting Factor	0.575	0.580	1%	-
% Correct Prediction	70.50%	71.20%	0.99%	-
		40 Alt SIS		
TI	0.04	0.63	*	0.59
Fitting Factor	0.538	0.545	1.39%	-
% Correct Prediction	66.7%	66.6%	-0.15%	-

*This figure was unrealistically high due to the low denominator (0.04). Hence, it was not reported.

Table 5.3 compares the performance of the full choice set with that of the 40 randomly and stratified importance-sampled alternatives. The analysis focused on the Base Utility model with and without the SC&AE parameter using the parameters estimated from the 2018 data applied to the 2022 data.

The Transfer Index (TI) revealed a significant impact of including the SC&AE parameter. For example, in the case of the randomly sampled choice set, the TI value without SC&AE was -1.21. This negative value indicates that the model performed worse than the locally estimated gravity model (simple reference model) when applied to the 2022 ODiN data. However, after including the SC&AE parameter, the TI improved to a positive value of 0.32. This improvement suggests that the model retained 32% of the performance gain that would have been achieved by re-estimating the Base Utility + SC&AE model using 2022 data.

Despite the significant improvement in the TI metric, improvements in other performance indicators are limited. The Fitting Factor increased by only 1%, and the percentage of correct predictions improved by only 0.99%. A similar pattern was observed when using the SIS and full choice sets. This discrepancy suggests that the high impact of the SC&AE parameter on the TI may be exaggerated. This limited impact is also observed in the prediction clearness analysis performed next for SIS (Figure 5.4).

This persistent trend of an exaggerated impact on the Transfer Index (TI) but a limited impact on other indicators is observed across all trip purposes. To explore the source of this exaggerated impact, Section 5.5.1 presents an analysis of the differences in log-likelihood (LL) values among the models used in the TI calculation using HBM trips with Random Sampling (RS) as an example.

Coming back to table 5.3, When using SIS, the TI value without the SC&AE parameter was already positive at 0.04, unlike the negative values observed with the full choice set and randomly sampled choice sets. This indicates that, with SIS, the model using the 2018 parameters performed better in 2022 than the gravity model estimated on the 2022 data, even without including the SC&AE parameter.

Interestingly, the percentage of correct predictions decreased slightly when using SIS. However, this metric has limitations because it does not account for the discriminatory ability of the model. Therefore, a small decrease in the percentage of correct predictions does not necessarily imply a negative impact on the model's performance. This is proven by the prediction clarity analysis, which consistently showed a positive impact below (Table 5.4).

Prediction Clearness

As the overall trend across the full and the two different sampling methods is the same, only the prediction clearness for SIS is analyzed in this section. Figure 5.3 illustrates the effect of SC&AE on the model's discriminative ability across thresholds that range from 40% to 90%. With the 40 alternatives chosen through stratified sampling, the impact of SC&AE remained consistently positive, but more limited in terms of % improvement, especially at thresholds below 80%. This is also evident in the zoomed-in version of the clarity analysis graph, which focuses on the "Clearly Right" predictions in figure 5.4, plotting the values in table 5.4

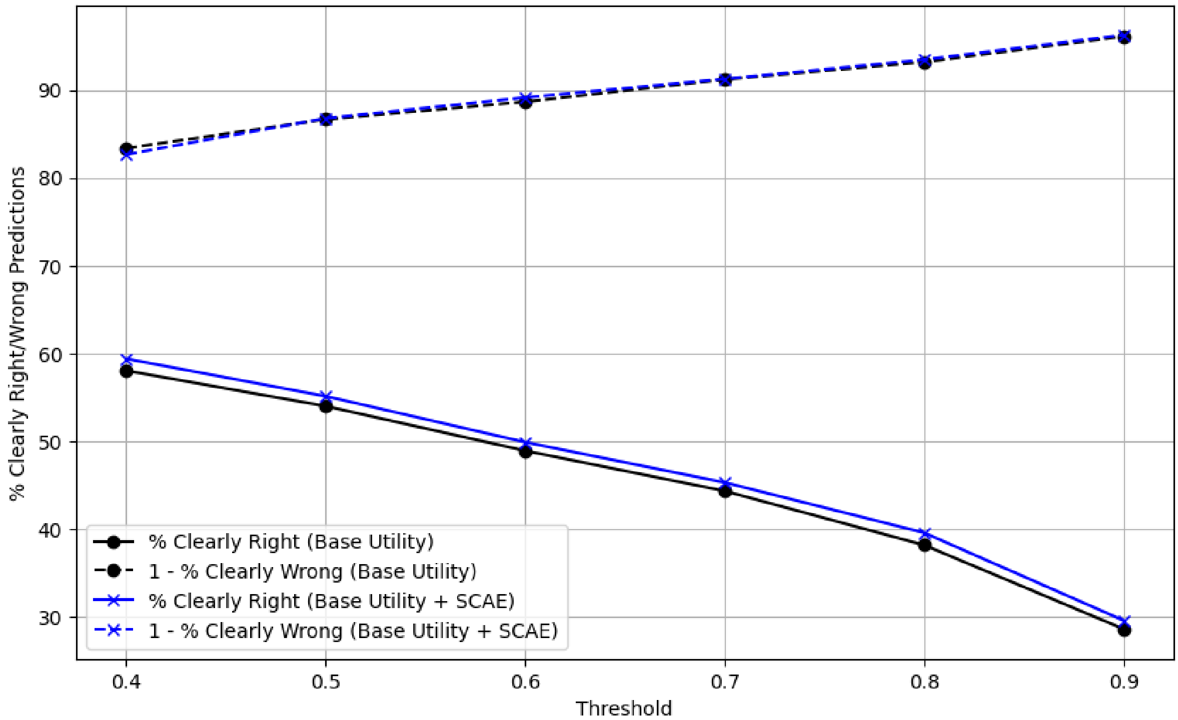


Figure 5.3: Clarity analysis graph for HBM Trips in 2022 Using 2018 Estimated Parameters (Stratified Importance Sampling)

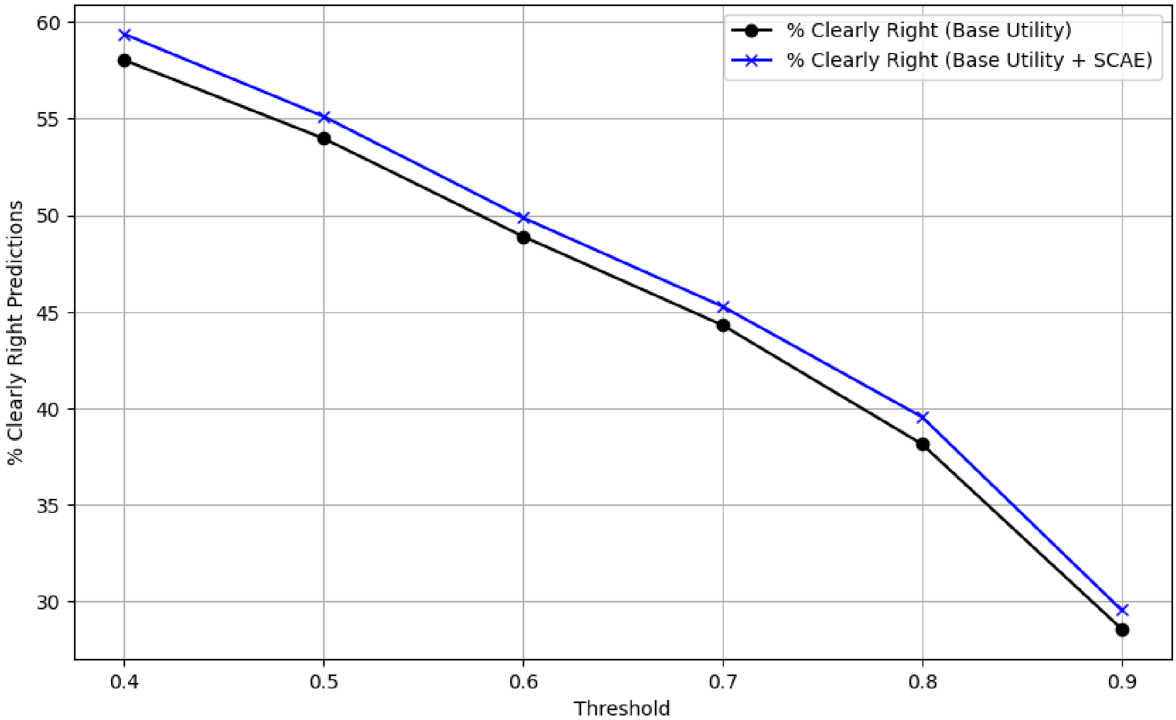


Figure 5.4: % Clearly Right for HBM Trips in 2022 Using 2018 Estimated Parameters (Stratified Importance Sampling)

Table 5.4: % Clearly Right values across thresholds for HBM (Base Utility vs Base Utility with SCAE, 40 alternatives (Stratified Importance Sampling)

% Clearly Right (HBM 2022), y18 parameters [40 SIS]			
Threshold	Base Utility	Base Utility+SCAE	% Improvement
40%	58.02%	59.37%	2.33%
50%	53.97%	55.11%	2.10%
60%	48.89%	49.86%	1.99%
70%	44.30%	45.27%	2.20%
80%	38.14%	39.55%	3.68%
90%	28.58%	29.55%	3.40%

5.2. Work Location Choice

A sampled choice set size of 45 destinations is used to perform the model estimation and then assess the impact of SC&AE on the temporal transferability of the destination choice model for work trips. Figures 5.5 and 5.6 illustrate the behavior of the statistically significant average beta values and their standard deviations across 30 iterations for stratified importance-sampled alternatives.

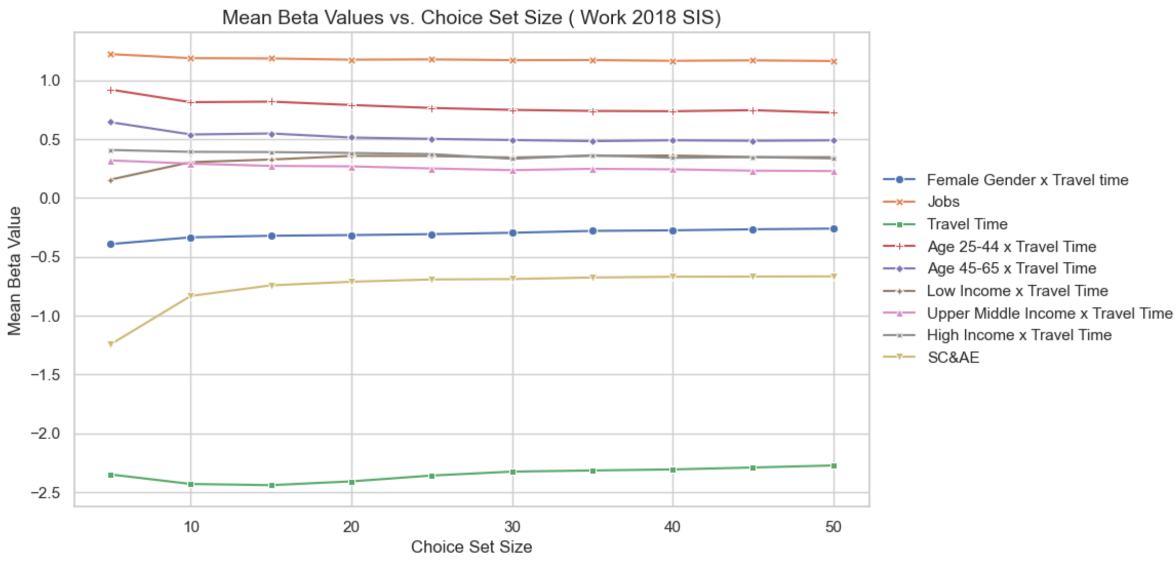


Figure 5.5: Mean Beta variation for Base Utility + SCAE with the choice set size (30 iterations/ choice set, Work 2018, Stratified Importance Sampling)

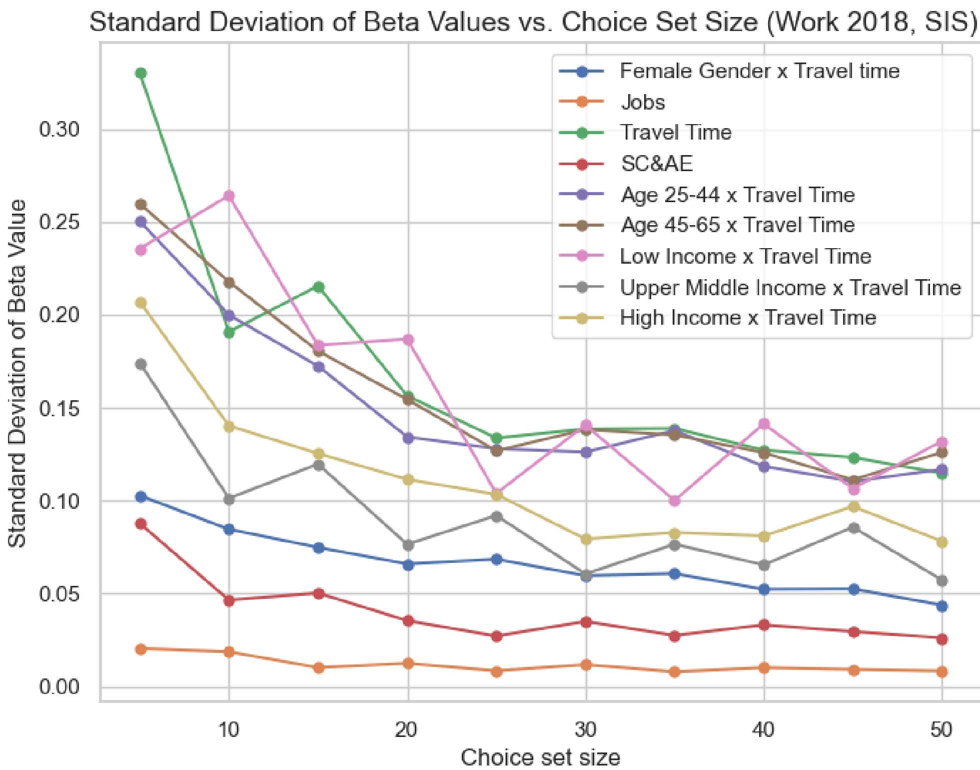


Figure 5.6: Standard Deviation of various parameters of Base Utility+SCAE with choice set size (30 iterations/ choice set, Work 2018, Stratified Importance Sampling)

As shown in Figures 5.5 and 5.6, in terms of standard deviations, the travel time is the least stable parameter. Hence, based on its stabilization, a choice set of 45 alternatives is identified as appropriate. Most other parameters exhibit their lowest standard deviations at 45 alternatives too.

5.2.1. Parameter Estimates

Table 5.5 presents the parameter estimates for both the full choice set and the sampled 45-alternative choice sets.

Table 5.5: Estimated Betas, Work (2018, 2022 ODiN), Full vs 45 Alt Choice set (Random and Stratified Importance Sampling)

Estimated Betas, Work (2018, 2022 ODiN), Full vs 45 Alt Choice set (RS and SIS)							
		2018			**2022**		
		Full	**45 RS**	**45 SIS**	**Full**	**45 RS**	**45 SIS**
Note	Name	Parameter Values (t-test values)			Parameter Values (t-test values)		
Male as reference level	Female Gender x Travel Time	-0.18 (-4.24)	-0.32 (-4.39)	-0.25 (-3.55)	-0.13 (-2.71)	-0.32 (-3.69)	-0.156 (-1.87)
Parameters for logarithmic values	Jobs	1.14 (45.4)	1.16 (42.4)	1.16 (42.66)	1.09 (36.3)	1.12 (34.1)	1.10 (34.27)
	SCAE	-0.68 (-6.79)	-0.52 (-5.48)	-0.67 (-7.09)	-0.71 (-7.14)	-0.68 (-6.29)	-0.767 (-7.15)
	Travel time	-2.01 (-20.6)	-2.21 (-11.5)	-2.35 (-12.1)	-1.99 (-19.2)	-2.60 (-12.2)	-2.317 (-12)
		Age Groups					
Age group 65+ as reference level	Age 25 - 44 x Travel time	0.57 (6.09)	0.74 (3.87)	0.68 (3.97)	0.46 (4.76)	0.90 (4.39)	0.69 (3.82)
	Age 45-65 x Travel time	0.38 (4.08)	0.55 (2.85)	0.48 (2.81)	0.30 (3.16)	0.73 (3.52)	0.45 (2.44)
		Income Groups					
Middle Income Group as reference level	Low Income x Travel time	0.323 (2.82)	0.10 (0.53)	0.428 (2.26)	0.299 (2.3)	0.49 (2.41)	0.39 (2.11)
	Upper Middle Income x Travel time	0.156 (2.62)	0.07 (0.67)	0.33 (2.98)	0.117 (1.74)	0.192 (1.60)	0.13 (1.07)
	High Income x Travel time	0.26 (3.66)	0.25 (2.01)	0.44 (3.46)	0.23 (2.73)	0.195 (1.24)	0.197 (1.37)
	rho square (null)	0.215	0.32	0.31	0.211	0.319	0.304

As shown in Table 5.5, more parameters are statistically significant at the 5% confidence level than those for HBM trips. Most parameters have the expected signs, except for the interaction between the low-income group and travel time.

Similar to HBM, we find a negative SC&AE parameter for Work, suggesting that spatial competition dominates the agglomeration effect in 2018 and 2022. These results are consistent with those of previous studies (Ho and Hensher, 2016).
Comparing the value of SC&AE parameter estimates for HBM and Work (SIS), HBM trips had a more negative parameter (-1.60 to -1.77) than Work (-0.67 to -0.76), indicating that SC&AE has a stronger influence on destination choices for HBM trips than for work trips. This seems reasonable from the perspective of ease of switching to alternative destinations. HBM trips, which include shopping and personal errands, often involve destinations that are closer substitutes (e.g., multiple grocery stores or service centers within a short distance), intensifying competition. On the other hand, work trips generally involve more specialized destinations (e.g., offices or job locations), where alternatives are more limited, resulting in weaker spatial competition. Additionally, travelers performing HBM trips have a high level of ease of switching because they have no mandate or commitment to stick to a specific shopping or service destination to perform maintenance activities. However, employment is a long-term decision with a longer commitment period, and switching jobs is not as easy as switching destinations to perform maintenance activities. Hence, the spatial distribution of opportunities has a lower influence on destination choices for work trips than for HBM trips.

Analyzing the interaction parameters reveals that the signs for income groups higher than middle income (upper-middle and high income) are as expected. Typically, as income increases, individuals tend to be less sensitive to travel time and are more willing to travel further; this trend was also observed by Bhat et al. (1998) for home-based work trips in the Boston Metropolitan Area, USA.

Given this reasoning, the positive sign and relatively high value for the low-income group is unexpected. This suggests that in the MRA, travelers in low-income groups are more willing to travel further than those in higher-income groups. Two factors may have played a role in this unexpected finding. One is the MRA's transportation system and the second is the spatial distribution of relevant employment opportunities.

First, unlike car-oriented developed urban areas, a typical feature of American cities such as the Boston Metropolitan Area, the MRA has a public transport-oriented infrastructure. This accessibility reduces reliance on private vehicles. In contrast, in car-oriented cities, the cost of owning and operating a vehicle may limit low-income travelers to shorter distances.
Second, the concentration and specialization of employment opportunities in the MRA may require greater travel distances for low-income workers. Jobs in specialized sectors or lower-income occupations are often located in areas distant from residential neighborhoods. Consequently, low-income individuals may need to travel further to access employment opportunities, even if this involves long commutes.

Comparing the values and statistical significance of the income parameters across both years reveals that while the high-income group parameter remains statistically significant in both years for the full choice set, the parameter for the upper-middle income group becomes statistically insignificant in 2022. This shift may indicate changing dynamics in travel behavior between 2018 and 2022.

Similar to HBM, the rho-null squared value in sampling increases here as well, over 48–50% in both years for work location choice, likely reflecting increased discriminatory power due to the limited choice alternatives in the sampled choice set.

5.2.2. Impact of SC&AE on temporal transferability

Transfer Index, Fitting Factor, and % Correct Predictions

Table 5.6: Work trips: Performance Comparison on various indicators Full vs. Sampled choice sets (RS and SIS)

Work 2022 (using 2018 estimated parameters)				
Full Choice Set				
Indicators	Base Utility	Base Utility + SCAE	% Improved	Absolute gain in TI value
TI	0.54	0.73	35.2%	0.19
Fitting Factor	0.0263	0.0272	3.42%	-
% Correct Prediction	7.93%	7.99%	0.75%	-
45 Alt Randomly Sampled				
TI	0.437	0.662	52.01%	0.22
Fitting Factor	0.147	0.152	3.61%	-
% Correct Prediction	26.06%	26.59%	2.03%	-
45 Alt SIS				
TI	0.48	0.71	47.91%	0.23
Fitting Factor	0.138	0.143	3.62%	-
% Correct Prediction	25.28%	25.59%	1.23%	-

Table 5.6 presents the results for 2022 work trips using the 2018 estimated parameters across full-choice, random sampling (RS), and stratified importance sampling (SIS) methods. Compared to HBM trips, including SC&AE in work trips exhibits smaller improvements in the Transfer Index (TI) and other indicators, reflecting weaker SC&AE impacts on temporal transferability for this trip purpose.

For the full-choice set, including SC&AE improves the TI from 0.54 to 0.73, representing a 35.2% increase. While this demonstrates a positive impact of SC&AE, the improvement is less dramatic than the jump observed for HBM trips (-0.44 to 0.82). Other indicators, such as the Fitting Factor and Percentage of Correct Predictions, also showed limited improvements (3.42% and 0.75%, respectively), further supporting the limited impact of SC&AE on the temporal transferability for work trips.

In the random sampling approach, TI increased from 0.437 to 0.662 with SC&AE, showing better transferability than that observed for HBM trips, where TI remained highly negative (-1.21 to 0.32).

With SIS sampling, the TI improves from 0.48 to 0.71, a 0.23 increase in absolute value, less than that for HBM using SIS (0.59). Notably, the Percentage of Correct Predictions shows a gain of 1.23% with SC&AE, in contrast to the slight decline (-0.15%) for HBM trips.

Interestingly, although a similar trend of exaggerated impact by TI and limited impact on other indicators is also observed here, unlike HBM trips where SC&AE had the most significant impact on TI, the improvements for work trips were more balanced across TI, Fitting Factor, and Correct Predictions. This suggests that, while SC&AE enhances transferability for work trips, its influence is less dominant than that of HBM trips. This is also reflected in the improvements due to SC&AE in % Clarity analysis presented next.

Prediction Clearness

in figure 5.8. For HBM trips (figure 5.4), the gap between the Base Utility and Base Utility + SC&AE specifications is more noticeable for HBM trips, indicating a stronger impact of SC&AE on HBM trips. By contrast, for work trips, the two curves become closer as the threshold increases.

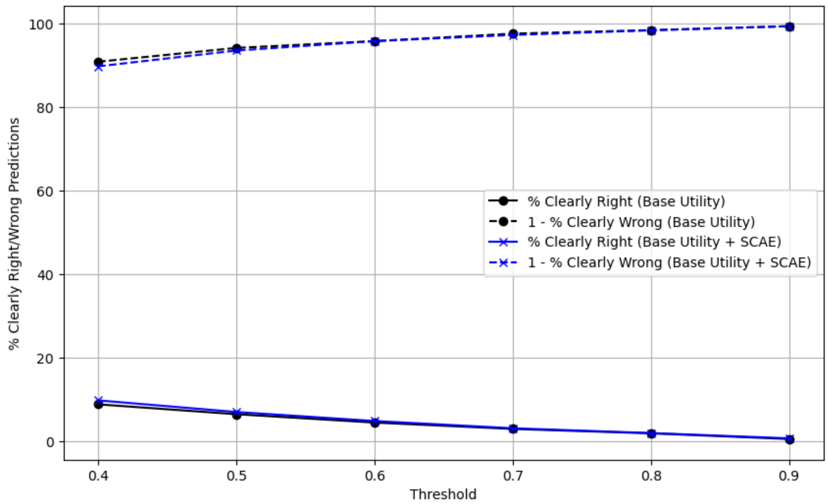


Figure 5.7: Clarity analysis graph for Work Trips in 2022 Using 2018 Estimated Parameters (Stratified Importance Sampling)

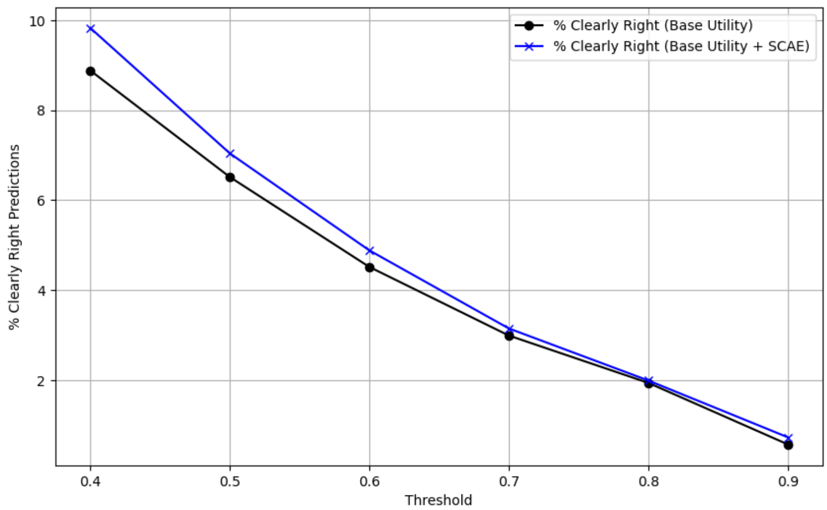


Figure 5.8: % Clearly Right graph, for Work Trips in 2022 Using 2018 Estimated Parameters (Stratified Importance Sampling)

Table 5.7: % Clearly Right values across thresholds, for Work Trips in 2022 Using 2018 Estimated Parameters (Stratified Importance Sampling)

% Clearly Right (Work 2022), y18 parameters [45 SIS]			
Threshold	Base Utility	Base Utility + SCAE	% Improvement
40%	8.88%	9.83%	10.65%
50%	6.52%	7.04%	8.06%
60%	4.52%	4.89%	8.14%
70%	3.00%	3.15%	5.26%
80%	1.94%	2.00%	2.70%
90%	0.58%	0.74%	27.27%

The values of %Clearly Right for Work across thresholds (0.74% to 9.83%) are far lower than those for HBM (29.55% to 59.37%), as shown in Table 5.4. This is because the work destination choice model has much lower range of rho-square values (0.301–0.304) than HBM trips (0.685–0.698). Naturally, the % improvement becomes higher due to the low values in % Clearly Right across thresholds. However, in terms of absolute change, the change in values due to SC&AE for Work is lower than that for HBM trips.

5.3. Secondary and higher education

For destination choices related to secondary and higher education trips, the results are presented using a 45-alternative choice set. Figures 5.9 and 5.10 illustrate the behavior of statistically significant average beta values and their standard deviations over 30 iterations for the Base Utility + SC&AE model specifications.

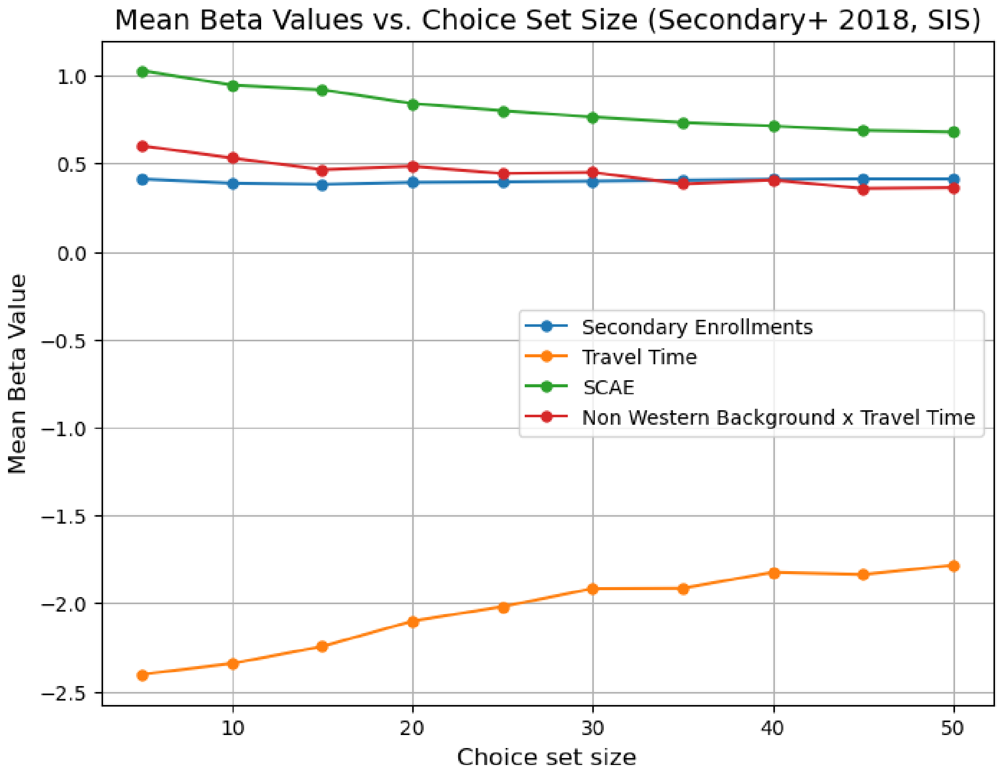


Figure 5.9: Mean Beta variation for Base Utility + SCAE with the choice set size (30 iterations/ choice set, Secondary+ 2018, Stratified Importance Sampling)

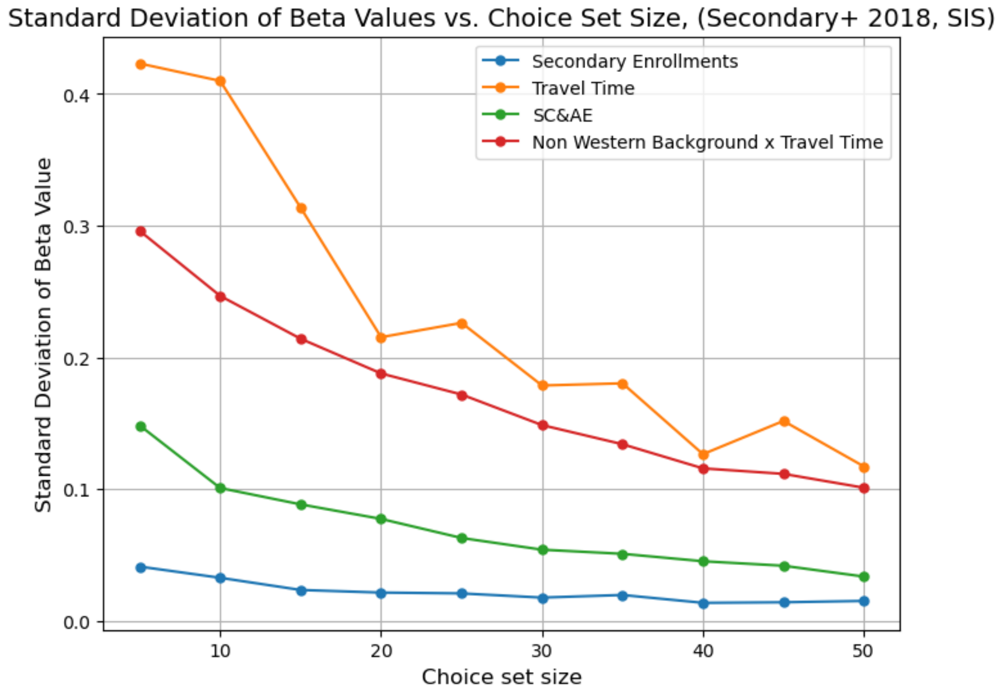


Figure 5.10: Standard Deviation variation of various parameters for Base Utility + SCAE with choice set size (30 iterations/ choice set, Secondary+ 2018 Stratified Importance Sampling)

As shown in Figures5.9 and 5.10, in terms of standard deviations, travel time is the least stable parameter. Ideally, 40 should have been the choice set size, but the standard deviation rose again for 45 alternatives. Hence, a choice set of 45 alternatives is identified as appropriate, as the standard deviation decreases from that point on.

5.3.1. Parameter Estimates

Table 5.8: Estimated Betas, Secondary and above (2018, 2022 ODiN), Full vs 45 Alt Choice set (Random and Stratified Importance Sampling)

Estimated Betas, Secondary and above (2018, 2022 ODiN), Full vs 45 Alt Choice set (RS and SIS)							
		2018			**2022**		
		Full	**45 RS**	**45 SIS**	**Full**	**45 RS**	**45 SIS**
Note	Name	Parameter Values (t-test values)			Parameter Values (t-test values)		
Parameters for logarithmic values	Secondary+ Enrollments	0.42 (7.14)	0.397 (6.6)	0.399 (6.822)	0.49 (6.08)	0.46 (5.6)	0.49 (6.28)
	SCAE	0.62 (4.33)	0.67 (4.55)	0.72 (4.61)	0.40 (2.07)	0.535 (2.73)	0.48 (2.51)
	Travel time	-1.70 (-13.8)	-1.74 (-8.44)	-1.97 (-7.77)	-1.78 (-11.7)	-1.94 (-8.71)	-1.89 (-7.13)
		Migration Background Groups					
Dutch as reference level	Non Western x Travel time	0.25 (2.11)	0.36 (2.07)	0.488 (2.68)	0.13 (0.79)	0.02 (0.08)	-0.12 (-0.46)
		Income Groups					
Middle Income Group as reference level	Low Income x Travel time	0.273 (1.62)	-0.01 (-0.03)	0.38 (1.31)	0.645 (3.35)	0.74 (2.73)	0.902 (2.94)
	rho square (null)	**0.227**	**0.277**	**0.273**	**0.222**	**0.28**	**0.276**

As shown in Table 5.8, only two interaction parameters related to traveler characteristics are statistically significant at the 5% confidence level. All parameters except SC&AE have expected signs.
Unlike the negative SC&AE parameter observed for the HBM and work trips, which indicates dominant spatial competition, SC&AE shows a positive parameter, signaling the dominance of agglomeration effects. This result contrasts with the findings of Sá et al. (2004), who reported a negative SC&AE parameter in their production-constrained gravity model for university choices among secondary school graduates. However, their model focused solely on one level post-secondary education choices, in which spatial competition may play a larger role. This study included trips for both secondary and multiple levels of higher education levels, Thus, students have a broader range of options for continuing education nearby in the choice set. This proximity to further educational opportunities can make zones with abundant secondary and post-secondary institutions more attractive, hence a positive value for dominating agglomeration effect.
Similar to work trips, the SC&AE and travel time parameters remain stable across both the full choice set and the 45-alternative sampled choice set. However, the interaction terms with travel time increase notably, with a 44 to 95% increase for the non-Western background group and a 14 to 40% increase for the low-income group by sampling methods.

As seen in the Work and HBM models, the rho-null squared value for the sampled choice set increases as well, rising by 22% to 27% in this case, indicating an enhanced model performance owing to the limited choice set.

5.3.2. Impact of SC&AE on temporal transferability

Transfer Index, Fitting Factor, and % Correct Predictions

Table 5.9: Secondary and Above Education trips: Performance Comparison on various indicators Full vs Sampled choice sets (RS and SIS)

Secondary and Above Education 2022 (using 2018 estimated parameters)				
		Full Choice Set		
Indicators	Base Utility	Base Utility + SCAE	% Improved	Absolute gain in TI value
TI	0.30	0.43	43.33%	0.13
Fitting Factor	0.0657	0.0632	-3.81%	-
% Correct Prediction	14.86%	12.68%	-14.67%	-
		45 Alt Randomly Sampled		
TI	-0.56	-0.05	91.07%	0.51
Fitting Factor	0.143	0.139	-2.88%	-
% Correct Prediction	33.70%	28.26%	-16.14%	-
		45 Alt SIS		
TI	-0.02	0.28	*	0.30
Fitting Factor	0.144	0.141	-1.81%	-
% Correct Prediction	32.25%	28.62%	-11.3%	-

*This figure was unrealistically high due to the low denominator (0.02). Hence, it was not reported.

Table 5.9 presents the results for secondary and above education trips in 2022 using the 2018 estimated parameters across the full-choice, random sampling (RS), and stratified importance sampling (SIS) methods. Overall, the

results on the performance indicators are inconsistent; positive on the TI metric but negative on the rest of the performance indicators. The source of this inconsistent performance is analyzed in section 5.5.2
Compared to the HBM and work trips, the improvements due to SC&AE for education trips are smaller on the TI metric and negative on other metrics. For the full-choice set, the inclusion of SC&AE improves TI from 0.3 to 0.43, a 43.3% increase, similar to work trips. However, the absolute gains are even lower. This reflects a weaker impact of SC&AE for secondary and above education trips. For random sampling, the TI improvement increases from -0.56 to -0.05, a good absolute increase of 0.51, yet it is still in the negative. It shows notable declines in the Fitting Factor (-2.88%) and the Percentage of Correct Predictions (-16.14%). This indicates that with random sampling, including SC&AE, the model performs poorly for secondary and above education trips. These results suggest that SC&AE has negligible relevance for secondary and above education trips.

With SIS sampling, including SC&AE, TI improves significantly from -0.02 to 0.28. This suggests that the impact of SC&AE on temporal transferability improves with a better sampling method, a trend observed in HBM and Work trips. This is also observed when comparing the performance on the % Clearly Right plot of Random Sampling (figure 5.11) and SIS (figure 5.12).

Prediction Clearness

To highlight the effect of sampling methods on the temporal transferability of destination choice models and, consequently, the impact of SC&AE on temporal transferability, this section compares the performance on the % Clearly Right plot of Random Sampling (figure 5.11) and SIS (figure 5.12).

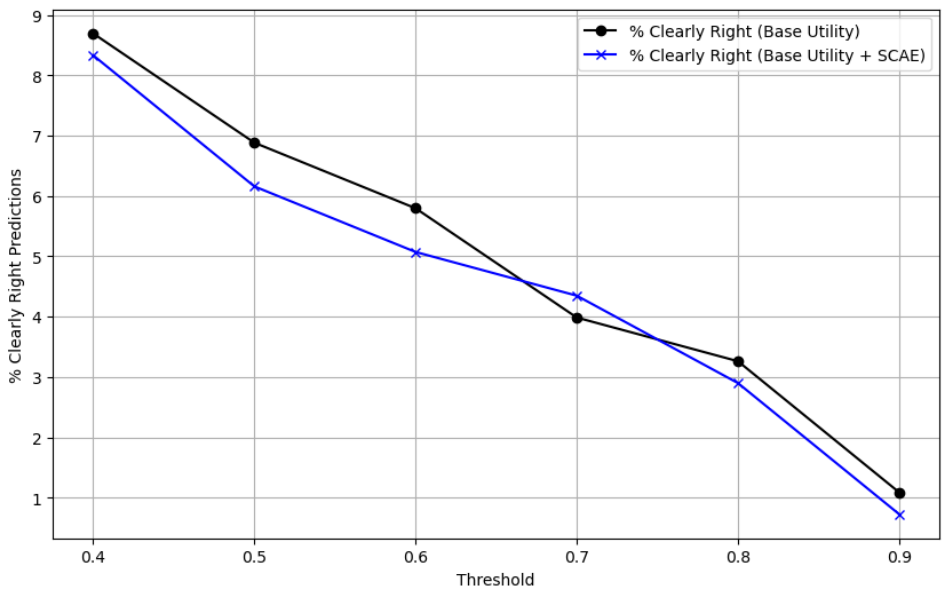


Figure 5.11: % Clearly Right graph, y18 parameters in 2022 (Base Utility vs Base Utility with SCAE, 45 alternatives, Random Sampling)

Table 5.10: % Clearly Right graph, for Secondary and above education trips in 2022 Using 2018 Estimated Parameters (Random Sampling)

% Clearly Right (Secondary+ 2022), y18 parameters [45 RS]			
Threshold	Base Utility	Base Utility+SCAE	% Improvement
40%	8.70%	8.33%	-4.17%
50%	6.88%	6.16%	-10.53%
60%	5.80%	5.07%	-12.50%
70%	3.99%	4.35%	9.09%
80%	3.26%	2.90%	-11.11%
90%	1.09%	0.72%	-33.33%

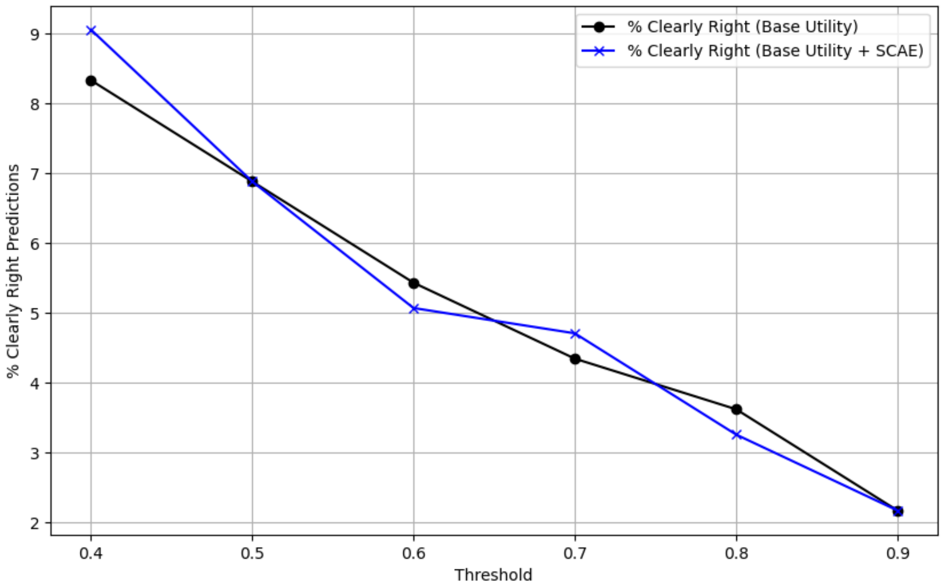


Figure 5.12: % Clearly Right graph, for Secondary and above education trips in 2022 Using 2018 Estimated Parameters (Stratified Importance Sampling)

Table 5.11: % Clearly Right performance comparison (Base Utility vs Base Utility with SCAE, 45 alternatives Stratified Importance Sampling)

% Clearly Right (Secondary+ 2022), y18 parameters [45 SIS]			
Threshold	Base Utility	Base Utility+SCAE	% Improvement
40%	8.33%	9.06%	8.70%
50%	6.88%	6.88%	0.00%
60%	5.43%	5.07%	-6.67%
70%	4.35%	4.71%	8.33%
80%	3.62%	3.26%	-10.00%
90%	2.17%	2.17%	0.00%

A comparison of the figures and tables above indicates that, with RS, the inclusion of SC&AE leads to a decrease in the Percentage of Clearly Right predictions across almost all thresholds. This suggests that RS may not effectively capture the benefits of accounting for SC&AE in destination choice models for secondary and above education level trips.
In contrast, SIS shows a relatively better performance with the inclusion of SC&AE, although the improvement is not particularly convincing. Several factors may explain this persistent underwhelming performance.

One, as mentioned when comparing performance on the rest of the indicators in the previous section, SC&AE might have negligible relevance for secondary and higher education trips.

Second, due to the low amount of observed data for each education level after primary education, this study combined travelers across different education stages into a single destination choice model for secondary and higher education trips. Such low numbers and aggregating travelers from various educational backgrounds fail to capture the essential differences in travel behavior among students at various stages of their education. Each educational stage comes with specific travel patterns. For example, the factors influencing destination choices for secondary school students can differ significantly from those affecting university students, because university students would have more autonomy and flexibility. Hence, betas will not be able to generalize well, leading to poor model performance in explaining destination choices. This limitation is reflected in the low rho-square values in Table 5.8 and consequently low percentages of observations assigned clearly right on the above % Clearly right table 5.11 across all thresholds.

While SIS demonstrates a better ability to capture the effects of SC&AE than RS, the underlying data limited in both quality and quantity hampers the model's performance. Consequently, it is not possible to draw definitive conclusions about the impact of SC&AE on the temporal transferability of the destination choice model for secondary and above education trips.

5.4. Primary education

5.4.1. Parameter Estimates

Table 5.12: Estimated Betas, Primary education (2018 ODiN), Full choice set

Estimated Betas, Primary education (2018 ODiN), Full choice set		
		2018
Note	**Name**	**Parameter value (robust t-test value)**
Parameters for logarithmic values	Primary enrolments	0.59 (7.97)
	SCAE	-0.22 (-1.23)
	Travel time	-1.85 (-20.81)
	Income Groups	
Middle Income group as reference level	Upper Middle Income x TT	-0.206 (-2.117)
rho square (null)		0.565

Table 5.12 presents the parameter estimates for primary education trips using the full choice set with the Base Utility + SC&AE specification. The three significant parameters have expected signs.
In the base year 2018, the parameter estimate for SC&AE is statistically insignificant, with a t-test value below the 1.96 threshold. Hence, the parameter estimation for 2022 and the rest of the analysis performed for the above trips are not performed. Although several parameters, including the SC&AE parameter in the Base Utility + SC&AE specification, were statistically insignificant, the model's rho square for 2018 remains notably high, at 0.565. This is the highest rho square achieved among all trip purposes considered in this study for the full choice set, with HBM trips as the next highest at 0.508, followed by secondary and higher education level trips at 0.227, and work trips at 0.215.
These results suggest that traditional modeling methods, which only incorporate travel time parameters and interaction effects, may effectively capture location choice behavior for primary education trips and that adding more nuanced parameters, such as SC&AE, might not be necessary.

5.5. Key Insights from Transfer Index performance Analysis

5.5.1. Source of Transfer Index's Exaggerated Impact

Table 5.13: TI Comparative analysis: Source of exaggerated Impact (HBM, Randomly Sampled)

	GM	Base Utility	Base Utility + SCAE
		2022	
Final LL	-2016.42	-2003.18	-1970.48
Gain over GM(2022)	—	13.24	45.94
		2022 (using 2018 parameters)	
Final LL		-2032.40	-2001.62
Gain over GM(2022)	—	-15.98	14.81
% Change in gain		-220.65%	-67.77%
TI		**-1.21**	**0.32**

Null LL:-6828.12

Table 5.13 shows the log-likelihoods of the models used in the Transfer Index calculation of the randomly sampled choice set for HBM. In the 2022 local estimation, the Base Utility specification shows only a modest gain of 13.24 in log-likelihood (LL) over the GM model. This limited gain likely stems from several traveler-related parameters being statistically insignificant at the 5% confidence level (t-test value < 1.96), as detailed in Table 5.3. In contrast, adding the SC&AE parameter yields a much larger gain of approximately 45.94 LL.

When using the 2018 estimated parameters, the Base Utility+SC&AE model still shows a gain of 14.81 LL, reflecting a 67.77% loss compared to what could have been achieved by re-estimating the model in 2022 (45.94). This explains the TI value of 0.32 i.e., it retains 32% of 45.94 LL, the gain over gravity model from re-estimation. Conversely, the Base Utility model, with 2018 parameters, performs slightly worse than the GM model, but the difference is small, only -15.98 LL. Although this difference is minor, it appears exaggerated compared to the minimal 13.24 LL gain achieved by re-estimating the Base Utility model, leading to a percentage change in the gain of -220.65%. This reflects how TI is calculated by comparing the loss in gains due to using base year parameter estimates to the potential gains from re-estimation (-15.98/13.24). Because the difference in LL between the Base Utility and GM models is small, focusing solely on TI performance can exaggerate the impact of the SC&AE parameter.

5.5.2. Source of Inconsistent Performance of SC&AE's impact for Secondary and above education level

Table 5.14: TI Comparative analysis: Source of Inconsistent Impact (Secondary and Above Education Location Choice, SIS)

	GM	Base Utility	Base Utility +SCAE
		2022	
Final LL	-770.79	-763.82	-760.24
Gain over GM(2022)	—	6.98	10.56
		2022 (Using 2018 parameters)	
Final LL		-770.91	-767.83
Gain over GM(2022)	—	-0.11	2.97
% Change in gain		-101.62%	-71.89%
TI		**-0.02**	**0.28**

Null LL: -1050.64

To explain why the inclusion of SC&AE has a positive impact on TI for secondary and above education trips, but fails to improve other performance measures, we need to compare the LLs of models used in the TI metric for HBM trips (Random Sampling) presented in the previous section 5.5.1 (Table 5.13) with those for secondary and above education trips (Table 5.14). For HBM trips, including SC&AE, achieves a TI value of 0.32, with absolute LL gains of 45.94 for re-estimation in 2022 and 14.81 using 2018 parameters over the reference GM model. In contrast, as seen in Table 5.14, for secondary education trips, the achievable TI value with SC&AE is similar at 0.28, but the absolute LL gains are much smaller: 10.56 from re-estimation and 2.97 using 2018 parameters over the GM model.

Clearly, the SC&AE parameter has a much stronger impact on improving model performance for HBM trips, as evidenced by the larger LL gains. However, because TI compares the ratios of these gains rather than their absolute values, it presents the maximum achievable transferability for SC&AE in both cases at similar levels. This explains why SC&AE had a positive impact on all indicators for HBM trips, yet showed a negative impact on these same indicators for secondary and above education trips, despite demonstrating a positive effect on the TI metric. The key issue lies in the relatively low absolute gains in the LL for secondary education trips and TI's reliance on the gain ratio to evaluate transferability.

5.5.3. Mitigation against parameter estimates errors due to sampling

Analyzing TI values for RS using Table 5.13 reveals that both the Base Utility and Base Utility+SC&AE specifications experience substantial losses in log-likelihood, resulting in lower TI values compared with the full choice set. The reason for this decline in TI performance can be better understood through a comparative analysis of the Base Utility and Base Utility+SC&AE parameter estimates presented in Tables 5.15 and 5.16. This comparison highlights the underlying cause of the decrease in the TI percentage and how including SC&AE mitigates the loss in TI value.

Table 5.15: HBM: Base utility parameter estimates in 2018 and 2022, Random Sampling

<table>
<tr><th colspan="8">Base utility: 40 RS</th></tr>
<tr><th colspan="2">LogLiklihoods</th><th colspan="4">Parameters</th><th>LL22 (y18)</th><th>TI</th></tr>
<tr><td colspan="6">2018</td><td rowspan="6">-2032.40</td><td rowspan="6">-1.21</td></tr>
<tr><td>LL Null</td><td>LL Final</td><td>Jobs</td><td>Travel Time</td><td>Lower Mid Income x TT</td><td>Upper Mid Income x TT</td></tr>
<tr><td>-6912.96</td><td>-2111.45</td><td>0.92 (24.82)</td><td>-3.20 (-17.84)</td><td>0.11 (0.48)</td><td>0.09 (0.48)</td></tr>
<tr><td colspan="6">2022</td></tr>
<tr><td>LLNull</td><td>LL Final</td><td>Jobs</td><td>Travel Time</td><td>Lower Mid Income x TT</td><td>Upper Mid Income x TT</td></tr>
<tr><td>-6828.12</td><td>-2003.18</td><td>0.82 (20.30)</td><td>-3.02 (-18.36)</td><td>-0.67 (-2.83)</td><td>-0.60 (-3.42)</td></tr>
</table>

Table 5.16: HBM: Base utility + SCAE parameter estimates in 2018 and 2022, Random Sampling

Base utility+SCAE: 40 RS								
LogLiklihoods		Parameters					LL22 (y18)	TI
2018							-2001.62	0.32
LL Null	LL Final	Jobs	Travel Time	Lower Mid Income x TT	Upper Mid Income x TT	SC&AE		
-6912.96	-2089.28	1.09 (22.92)	-3.10 (-17.36)	0.05 (0.22)	0.09 (0.49)	-1.35 (-7.17)		
2022								
LLNull	LL Final	Jobs	Travel Time	Lower Mid Income x TT	Upper Mid Income x TT	SC&AE		
-6828.12	-1970.49	1.02 (19.73)	-2.96 (-18.40)	-0.73 (-3.05)	-0.57 (-3.30)	-1.67 (-9.20)		

As shown in Tables 5.15 and 5.16, the interaction between income groups and travel time for lower and upper middle income groups in the base utility specification, which was statistically insignificant in 2018 at the 5% confidence level, becomes significant in 2022. This results in a high loss of explanatory power when the 2018 parameters are applied to the 2022 ODiN data. In the full choice set, the model only experienced a loss of 3.13 LL, but in this case, it lost more than five times by approximately 16 LL. Compared to the potential gain from re-estimating the model (13.24 LL), the losses increase by approximately 77% (from 143.74% in the full choice set to 220.65%).
The base utility+SC&AE specification faces the same problem. However, the losses increase by only approximately 50% (from 18.25% to 67.77%). This instability in the statistical significance of the income group interaction parameters may be due to sampling errors, as the Upper Middle-income parameter was statistically significant. The lower-income parameters were insignificant for 2018 and 2022 in the full choice set, as shown in Table 5.3. However, including the SC&AE parameter reduces the loss in the percentage change in gain by approximately 27%.

5.6. Discussion

This study is the first to validate the theory of Spatial Competition and Agglomeration Effects (SC&AE) beyond goodness-of-fit measures typically performed for a single period. Previous studies have demonstrated that including SC&AE parameter in destination choice models improves model fit and explanatory power, typically finding a negative sign for the SC&AE parameter, indicating a dominant spatial competition effect. The findings of this study are consistent with previous research, showing negative SC&AE parameters for home-based maintenance (HBM) and work trips in both 2018 and 2022.

First, regarding the findings from the Results chapter which are consistent with previous research, HBM trips have a more negative SC&AE parameter (-1.60 to -1.77) than Work trips (-0.67 to -0.76), indicating a stronger influence on destination choices for HBM trips. This is because HBM trips, such as shopping and errands, involve easily substitutable destinations with high ease of switching (e.g., multiple nearby stores) and no mandate/commitment for travelers towards a destination, intensifying spatial competition. By contrast, work trips involve specialized destinations with fewer alternatives and longer-term commitments, making switching less feasible. Therefore, the spatial distribution of opportunities has a greater impact on HBM trip destinations than on Work trips.

For secondary and higher education trips, the positive SC&AE parameter suggests that agglomeration effects dominate spatial competition. Unlike Sá et al. (2004), who found dominating spatial competition in university choices among high school graduates in the Netherlands by focusing solely on one post-secondary education level, this study included both secondary and multiple higher education levels. This broader range of nearby educational options to continue education makes areas with a higher number of institutions more attractive, resulting in a positive value that reflects dominant agglomeration effects. For primary education trips, the SC&AE parameter in 2018 was statistically insignificant, indicating a negligible influence on destination choices. This could be due to young children having the highest commitment period and the lowest flexibility to switch schools. Hence, the spatial distribution of primary school opportunities has little effect on their location choices.

Addtionally, this research extends beyond evaluating SC&AE's statistical significance in a single context by exploring its contribution to the temporal transferability of destination choice models, i.e., their ability to maintain predictive accuracy in subsequent forecasting years.
The findings indicate that SC&AE has a positive but limited impact on the temporal transferability of destination choice models. This impact varies by trip purpose, being highest for HBM trips, followed by work trips, and inconsistent for secondary and higher education trips (positive on the Transfer Index but negative on rest of the indicators).

There are some interesting findings from these results. Notably, the Transfer Index (TI) performance metric tends to exaggerate SC&AE's positive impact when viewed in isolation; a closer examination of log-likelihoods and other performance indicators (such as Fitting factor, % of correct predictions, and Clearness of predictions) confirms that the actual impact is limited. This exaggeration occurs because TI relies solely on the ratio of gains in log-likelihoods (LL), which can misrepresent models with small absolute gains. For example, a model with an LL gain ratio $\frac{1}{2}$ and another with a gain ratio of $\frac{50}{100}$ will have the same TI value of 0.5. But clearly, the second model is much better and will perform positively on other indicators, while the model with LL gain ratio $\frac{1}{2}$ will perform poorly on other indicators. Such a case was also discussed in Section 5.5.2.
Therefore, TI values should always be presented alongside other performance measures or at least be accompanied by a comparative analysis of the log-likelihood values used in their calculation.

Observing the trend of varying impacts of SC&AE on temporal transferability by trip purpose, the impact of SC&AE decreases with decreasing autonomy and ease of switching. For HBM trips, travelers have the highest autonomy and ease of switching. Shoppers have no mandate or commitment to stick to a specific shopping or service destination to perform maintenance activities. Work trips come second in terms of the ease of switching to alternate destinations. While travelers have a high level of autonomy because only travelers of legal working age are considered (excluding the 6–14 age group), the ease of switching destinations is lower than for HBM trips. Employment is a long-term decision with a higher commitment period, and switching jobs is not as easy as switching destinations to perform maintenance activities. However, there is still some flexibility as individuals can choose their commitment period based on personal preferences and job market conditions.

For education trips, the number of trips post-filtering is relatively low compared to work and HBM trips. Consequently, the parameters have less room for statistical significance (i.e., t-test values above 1.96). Moreover, because of the relatively low number of observations for each attained education level for travelers, travelers with different attained education levels were combined for secondary and higher education trips. Hence, the destination choice model developed for this trip purpose fails to capture the important differences in travel behavior among students at various stages, undermining the results. Therefore, the impact of SC&AE on the temporal transferability of education trips should be considered inconclusive. However, considering autonomy and ease of switching, the impact is likely to be limited, more so than for the HBM and work trips. This is because travelers (mostly likely students) must continue attending the same educational institution until they complete the required years or criteria for attaining a certain degree, effectively "locking them in" for a fixed period, irrespective of their preference, unless in exceptional circumstances.
Lastly, for primary education trips, autonomy and flexibility are the lowest, as all travelers are young children and primary education is mandatory. Therefore, they are committed to a fixed educational path that is even longer than secondary education. Hence, the spatial distribution of primary school opportunities has little effect on their location choices, resulting in insignificant SC&AE parameter.

Additionally, the choice of sampling method affects the impact of SC&AE on temporal transferability. Models using Stratified Importance Sampling (SIS) show higher TI values than those using Random Sampling (RS), with and without the SC&AE parameter. While the initial performance boost from SIS reduces the absolute gain in the TI value from including SC&AE compared to RS, SIS allows models to achieve higher overall TI values after including SC&AE.
Moreover, SC&AE offers protection against errors in parameter estimates due to sampling, as seen in section 5.5.3, where when randomly sampled, income group parameters turned insignificant in 2022 for HBM and cause a high loss in TI at a % change in log-likelihood gain of -220.65%. Even though the same phenomenon occurred even

after including SC&AE, the% change in the log-likelihood gain was limited to only -67.77%. Since sampling is almost necessary when dealing with a high number of destinations, this feature of SC&AE is particularly valuable, and it might not have been revealed if only its impact on explanatory power in a single year had been studied.

Overall, SC&AE had a limited but positive impact on the model's temporal transferability. Given the minimal effort required to include these effects in an MNL model because it reuses existing information such as zonal size measure and travel impedance, SC&AE provides technically "free" robust log-likelihood gains, especially for trips where travelers have high autonomy and low cost of switching destinations.

(Implications due to methodological choices in data processing are discussed in Section 4.5 in order to keep this section concise. Additionally, there is an error in ODiN Data Processing assumption for HBM trips. Please refer to the footnote at the end of that section)

From a policy perspective, the results suggest that including SC&AE in destination choice models enhances their effectiveness as predictive tools by improving temporal transferability. Models that include SC&AE maintain predictive accuracy over time and are particularly valuable for scenarios in which travelers have significant autonomy and flexibility, such as discretionary activities such as shopping and maintenance trips. This added robustness stems from SC&AE's ability to address two fundamental flaws of Multinomial Logit (MNL) models that limit their behavioral accuracy in representing travelers' destination choices: (1) the Independence of Irrelevant Alternatives (IIA) assumption, and (2) neglecting the influence of the spatial distribution of opportunities. By including information about all alternative destinations, SC&AE attempts to tackle both issues, allowing MNL models to become more behaviorally representative while retaining computational simplicity.

The results of this study have broader implications for transport modeling. Traditionally, to overcome the limitations of Multinomial Logit (MNL) models, researchers have relied on more complex disaggregate models that are computationally intensive and often impractical for large datasets. However, this study demonstrates that simpler models such as MNL can overcome their flaws by including behavioral theories such as Spatial Competition and Agglomeration Effects (SC&AE), thereby improving behavioral representation while retaining computational efficiency.

By focusing solely on whether a model explains or predicts behavior well, we may have been asking the wrong questions. Instead, we should ask, "Is my model an accurate representation of the system it is supposed to represent?". In transport modelling, this usually pertains to an accurate behavioral representation. By addressing this fundamental question and bringing models closer to accurately reflecting the system, we automatically enhance their explanatory and predictive capabilities applicable to the context. For example, Stratified Importance Sampling (SIS) in destination choice models improved the temporal transferability for all trips in this study, even without SC&AE, when compared to Random Sampling (RS). This is likely because SIS better represents how travelers might form their choice sets than RS does.

The findings from this study show that we do not necessarily need to rely solely on more complex models; there is another way: using theories to enhance simpler models. To achieve this, we need to look beyond the transport domain and draw insights from related fields, such as psychology or other behavioral sciences. It may be time for transport modelers to use relevant psychological theories to make transport models better representations of how travelers make choices. This approach does not mean we should focus only on simple models; at a certain point, these models will reach their limit in how much they can be improved using behavioral theories. Rather, this approach offers another pathway to improve transport models, allowing us as transport modelers to make complex models computationally feasible and use data more efficiently rather than relying primarily on advancements in computational science. The results from this study are certainly encouraging.

6
Conclusion, Limitations, and Recommendations

6.1. Conclusion

This study aimed to assess the validity of SC&AE in explaining destination choices more holistically, thereby justifying its inclusion in destination choice models (DCMs). It uses a case study of the Metropolitan Region of Amsterdam (MRA), focusing on work, education, and home-based maintenance trips arriving in the MRA in 2018 and 2022. To achieve this, the study goes beyond traditional goodness-of-fit statistics by evaluating the contribution of SC&AE to temporal transferability, thus taking a step toward a more comprehensive approach for validating SC&AE. Specifically, it compares the performance of a model specification excluding SC&AE effects with one that includes them, using various performance indicators such as the Transfer Index, Fitting Factor, % Correct predictions, and Prediction Clearness to assess the models.

Accordingly, the main research question was formulated as follows:

How do SC&AE affect the temporal transferability of destination choice models ?

As analyzed in Chapter 5 Results and subsequently discussed in Section 5.6, SC&AE has a positive but limited impact on the temporal transferability of destination choice models. This impact varies by trip purpose, with the highest being for HBM trips, followed by work, and inconsistent for secondary and above education levels. For primary education, the SC&AE parameter was found to be statistically insignificant in 2018; thus, the analysis was cut short at this point for primary education trips.
Observing this trend of varying impacts of SC&AE on temporal transferability by trip purpose, when considering autonomy and the ease of switching to alternative destinations, the impact of SC&AE on temporal transferability decreases with decreasing traveler autonomy and ease of switching destinations; it is highest for HBM trips, where travellers have high autonomy and flexibility, less so for work trips due to longer commitment periods, and inconsistent (positive on the Transfer Index but negative on rest of the indicators) for secondary and above level education trips, where travellers are effectively committed to institutions until they complete their education. For primary education trips, autonomy and flexibility are the lowest, thus explaining the statistically insignificant estimated SC&AE parameter.

Four sub-questions were formulated to answer the main research question. The methodology outlined in Figure 4.1 consists of phases of Exploration, Design, and Validation, which are applied to address the sub-questions in a structured manner. The sub-questions and their respective answers are as follows:

1. What factors affect the temporal transferability of destination choice models? (Section 2.4, Chapter 5)

   From previous research, the key factors affecting the temporal transferability of destination choice models include **Regional characteristics** (e.g., infrastructure changes and population growth) that introduce new behaviors not present in the previous context and, hence, not captured in the original model, **Past travel behavior (inertia)**, where stable patterns maintain the model's relevance over time. **Forecasting horizon** is also expected to impact transferability, as longer horizons increase exposure to significant behavior changes, though the evidence is inconclusive. **Overfitting** reduces transferability, as models overly reliant on specific contexts lose predictive power with changing conditions. **Model specification** improves stability when using more socioeconomic characteristics of travellers to explain choices, the **IIA assumption** of MNL models potentially reduces transferability by leading to unrealistic predictions in the presence of correlated alternatives, although the findings were inconclusive.

Moreover, the findings from this study (Chapter 5) suggest that the **Choice of Sampling Method** also affects the temporal transferability of the destination choice models. Models using Stratified Importance Sampling (SIS) show higher TI values than those using Random Sampling (RS) with and without the SC&AE parameter. Overall, the findings indicate that improving the **Behavioral representation** of destination choice models positively impacts their temporal transferability.

2. How do existing destination choice models incorporate Spatial Competition & Agglomeration Effects (SC&AE), and what is the significance of these effects on destination choices? (Section 2.3)

Existing destination choice models include SC&AE primarily through the Hansen-type accessibility index which includes travel impedance and varies the attraction factor based on the trip purpose. For example, total jobs for work trips, service and retail employment for home-based maintenance (HBM) trips, and total student enrollment for educational trips. A negative estimated value of the index indicates that spatial competition dominates, whereas a positive value suggests that agglomeration effects are more dominant. If the parameter value is zero, it indicates that both are equally strong or that both are absent. Some studies segregate this net effect-capturing single index into two separate accessibility factors: one for complements (different industries/sectors) and another for substitutes (similar industries/sectors). This approach helps identify the sources of spatial competition and agglomeration, which vary according to the trip purpose. Previous research finds that competition forces arise from complements for work purposes and from substitutes for HBM and HBO purposes.
Across all trip purposes, including SC&AE improves model fit, with spatial competition effects (indicated by negative parameters) generally dominating, although the intensity varies; for instance, spatial competition is stronger for Home-based other (HBO) trips than for HBM trips. For all trip purposes, spatial competition effects dominate (negative parameter), albeit varying in intensity and source. For example, in one study, spatial competition was stronger for HBO trips than for HBM trips. Overall, all studies show a significant SC&AE parameter and improvement in goodness of fit owing to the relaxation of the IIA assumption.

After developing the utility functions (Section 4.6) for each trip purpose by selecting variables (Section 4.2) based on the findings from the Exploration phase, the results in the validation phase answers the third subquestion:

3. How do SC&AE impact the destination choices of travellers in the study area, and how do they differ across trip purposes? (Chapter 5 Results, Section 5.6)

SC&AE influences destination choices differently depending on the trip purpose within the study area. For home-based maintenance (HBM) and work trips, the SC&AE parameter is negative, indicating that spatial competition dominates. This effect is stronger for HBM trips (with values between -1.6 and -1.77) than for work trips (-0.67 to -0.76). This difference likely arises because work trips involve long-term commitments, which reduce destination flexibility.
For secondary and above education level trips, agglomeration effects dominate (positive parameters). This study included both secondary and above education levels. Hence, the broad range of nearby educational options to continue education makes areas with a higher number of institutions more attractive. For primary education trips, the estimated SC&AE parameter was statistically insignificant in 2018; therefore, no further analysis was conducted. The statistically insignificant parameter indicates the negligible influence of SC&AE on destination choices. This could be because young children have the highest commitment period and the lowest flexibility to switch schools. Hence, the spatial distribution of primary school opportunities has little effect on their destination choices.

4. How do the DCMs for various trip purposes perform across statistical tests and predictive measures, and what are their policy implications for the study area? (Chapter 5 Results, section 5.6)

The findings indicate that including SC&AE in destination choice models has a positive but limited impact on their temporal transferability, varying by trip purpose, with the highest impact for home-based maintenance (HBM) trips, followed by work trips, inconsistent for secondary and higher education trips, and a statistically insignificant parameter for primary education trips in 2018. Thus, the impact of SC&AE on temporal transferability decreases with decreasing traveler autonomy and ease of switching destinations, higher autonomy and flexibility (as in HBM trips) lead to a higher SC&AE impact, while lower autonomy and flexibility (as in primary education trips) results in statistically insignificant parameter. However, due to the low amount of observed data for each education level after primary education, this study combined travellers across different education stages into a single destination choice model for secondary and higher education trips. Thus, these factors undermined the results of education trips. Therefore, the impact of SC&AE on the temporal transferability of education trips should be considered inconclusive. However, considering autonomy and ease of switching, the impact is likely to be limited, more so than for the HBM and work trips.

Regarding the performance on various indicators, the Transfer Index metric exaggerates SC&AE's positive impact when viewed in isolation. The log-likelihoods of the models used in the TI calculation and other indicators such as Fitting factor, % Correct prediction, and Discriminative ability show its actual, limited effect. This exaggeration occurs because TI relies solely on the ratio of gains in log-likelihoods (LL), which can misrepresent models with small absolute gains.
Thus, TI values should always be presented with additional performance measures or at least the log likelihood of models in the TI calculation.

Moreover, the choice of sampling method affects the temporal transferability of destination choice models, and thus, the impact of SC&AE on temporal transferability. Models using Stratified Importance Sampling (SIS) show higher TI values than those using Random Sampling (RS) with and without the SC&AE parameter. While the initial performance boost from SIS reduces the absolute gain in the TI value from including SC&AE compared with RS, SIS allows models to achieve higher overall TI values after including SC&AE. Thus, including SC&AE in Multinomial Logit (MNL) models provides "free" robust log-likelihood gains with minimal effort (especially for trips where travellers have high autonomy and low destination switching costs), as it reuses existing data like zonal size measures and travel impedance.

From a policy perspective, including SC&AE enhances the temporal transferability of destination choice models, maintaining predictive accuracy over time, particularly valuable for discretionary activities, such as shopping and maintenance trips. By including information about all alternative destinations, SC&AE addresses two fundamental flaws of MNL destination choice models: the Independence of Irrelevant Alternatives (IIA) assumption and negligence of the influence of the spatial distribution of opportunities. This allows MNL models to become more behaviorally representative without sacrificing their computational simplicity.
From a broader perspective, to improve transport models in general, we need to focus on improving their behavioral representation. To do so, we need not rely solely on complex models; we can enhance simpler models using theories. Transport modelers can better represent how travellers make choices using insights from other related fields, such as psychology, to develop their own theories. When these simple models reach their limit on how much they can be improved using behavioral theories, which they eventually will, this approach might allow another pathway to improve more complex models, making them more computationally feasible and data-efficient without relying primarily on advancements in computational science.

6.2. Limitations

This study's forecasting horizon is limited to five years (2018-2022) due to data availability constraints with the ODiN travel survey. Destination choices tend to evolve gradually and hence, policymakers are often interested in long-term forecasts (seven to ten years or more) to plan infrastructure, develop policies, and effectively anticipate future transportation needs. Without longer forecasting horizons, this study may miss gradual shifts in travel behavior, such as changes in residential patterns, employment locations, or the development of new urban centers. These shifts are crucial for understanding the actual impact of SC&AE on destination choice models as predictive tools. A longer forecasting horizon would have allowed for a more robust analysis of how SC&AE parameters perform over time, and how this theory could enhance the model's applicability as a predictive tool for long-term planning.

The chosen study period overlapped with the COVID-19 pandemic, a high-impact global event that significantly disrupted travel behaviors due to lockdowns, social distancing measures, and shifts toward remote work and online education. The findings may pertain to potentially altered behaviors due to the pandemic, and not patterns that would have emerged under a business-as-usual scenario not marked by such a rare and significant event. As a result, the findings could reflect temporary behavioral changes rather than stable trends, limiting the generalizability of the results to business-as-usual contexts. The absence of a scenario analysis comparing business-as-usual and pandemic-affected scenarios leads to the inability to isolate the effects of SC&AE from those induced by the pandemic, thus limiting our understanding of the true impact of SC&AE in various circumstances.

Furthermore, the travel time matrix in this study was based on the mode most frequently used in the survey for each trip purpose, which, while practical, overlooks the impact of other available modes. Although a weighted average approach was tested to account for mode availability (Section 3.4.2), it resulted in unrealistically high travel times. Relying on a single mode for travel time estimation overlooks the influence of other available modes on the destination choice. This simplification can lead to inaccuracies in modeling travel impedance, which is a key determinant of travellers' destination decisions. The model fails to capture the actual accessibility of destinations by not adequately accounting for the availability of various modes. Hence, this limitation should be considered when interpreting the results, as SC&AE and other model parameters interact with travel time for all trip purposes.

Due to the low amount of observed data for each education level after primary education, this study combined travellers across different education stages into a single destination choice model for secondary and higher education trips. Ideally, richer data regarding traveler trip observations and information on the institutions, such as education level and course offered, should be used, as these factors play a major role in explaining destination choices for education trips, as seen in previous research. Consequently, using this richer data, separate models for each education level should have been developed to capture the unique factors influencing destination choices specific to education levels. This limitation fails to capture the essential differences in travel behavior among students at various stages, thereby undermining the results. Each educational stage comes with specific travel patterns. For example, the factors influencing destination choices for secondary school students can differ significantly from those affecting university students because university students would have more autonomy and flexibility. These factors, i.e., low number of observations, absence of information regarding the quality of opportunity, and aggregating travellers that might showcase different patterns for education trips, render findings on the impact of SC&AE on temporal transferability as inconclusive for secondary and above education trips.

Finally, the weighting factor was not considered in this study for the ODiN travel survey data due to time constraints. The weighting factor in the ODiN survey is an adjustment made to the survey data to ensure that the sample accurately represents the broader population in the Netherlands aged six and above by compensating for the response bias, especially among survey respondents who are on holidays, and adjusting for sample selectivity (Statistics Netherlands (CBS), 2024). Underrepresented or overrepresented demographic groups in the unweighted sample may disproportionately influence the findings, leading to inaccurate population-level inferences. Thus, not applying ODiN's weighting factors limits the representativeness of the study, as it excludes adjustments for demographic, seasonal, and holiday-related biases. This omission may skew the results, making population-level inferences less accurate because the findings may reflect only sample-level trends rather than accurate estimates for the entire population.

6.3. Recommendations and Future research:

The findings suggest that including SC&AE into an MNL destination choice model requires minimal effort, as it leverages existing information such as zonal size measures and travel times, providing technically "free" robust log-likelihood gains. Therefore, the main recommendation is to first account for the limitations mentioned above and perform the required calibrations. and then include SC&AE parameters in destination choice models, especially for trips where travellers have significant autonomy and flexibility in choosing destinations, such as shopping and maintenance activities.

Although not directly related to SC&AE, an interesting finding is the method for comparing different sampling techniques and their effectiveness in parameter estimation. In this study, it was computationally feasible to include all zones in the analysis and to compare the estimates from different sampling methods with the full choice set. In situations where the total number of zones is too high for computational feasibility, researchers can aggregate these zones into smaller numbers, allowing for parameter estimation across all aggregated zones. They can then compare this aggregated full-choice set with different sampling methods in

terms of parameter estimation, similar to the approach used in this study. This method serves as a proxy for identifying a better sampling technique, at least in terms of its proximity to the full choice set results. However, caution is needed, as some parameter values may be distorted or vary significantly, given that specific effects can depend on the zonal size. Hence, this method is a modeling choice, and modelers must consider this, along with other influencing factors, when using aggregated zones to compare sampling methods.
The validation phase methodology to provide quantitative evidence of the impact of SC&AE on the temporal transferability of DCMs can be further adapted into a more comprehensive experimental scenario analysis framework to test any other theory, enhancement, or comparison of the different choice models as described in the example below using Table 6.1

Table 6.1: Example of Experimental framework for comprehensive scenario analysis

	Scenarios				
	Business As Usual			Global Impact Event	Other Scenarios
	Short term (Upto 5 years)	Medium (7 to 15 years)	Long (15 to 20 years)	Various Forecast horizons	Various Forecast horizons
Base Utility					
Base Utility + SC&AE					
Base Utility + Theory A					
:					
:					
N Specifications					

Should a high-impact pandemic-like event occur in the future, the insights gained from this study enable researchers to make better-educated guesses about the model's performance in such scenarios. They can experiment with different specifications, integrate theories beyond the SC&AE in the utility function, and assess the performance of these alternative specifications during short-term, high-impact global events. This flexibility allows them to identify the most robust specification for building an effective destination choice model. Furthermore, as additional ODiN data becomes available in subsequent years, for example, until 2030, researchers can adjust the forecasting horizon to examine the model's transferability over different forecasting horizons. For instance, using the earliest available data from 2018, they can project medium-term scenarios (5-10 years) using data from 2028, and perform a thorough comparative analysis of various specifications. In the absence of significant disruptive events in these subsequent years, the model's performance in a normal scenario can also be tested, enabling a comprehensive scenario analysis to build confidence in the model's forecasting capabilities among transport planners and other stakeholders.

For future research, it is recommended to address mode assumptions by using travel times that include all available transportation modes. Exploring the use of the mode log sum appears promising in large transport models, where mode choice is assumed after destination choices, as it captures the combined accessibility effects of different modes and reflects travellers' mode choices more accurately. Adapting the SC&AE formulation to integrate the mode log sum could enhance the model's representation of accessibility, providing a more comprehensive understanding of how various transportation options influence destination choice.
Additionally, this framework can be utilized to assess the impacts of various enhancements on subsequent choice models within a comprehensive transport modeling system, which often includes multiple interconnected models such as route choice, mode choice, and route assignment. By examining how improvements in the destination choice model, such as including SC&AE, influence other components, we can better understand the cascading effects throughout the modeling framework and how they contribute to overall predictive performance.

Further research could also investigate more complex formulations of SC&AE, such as defining them using two accessibility factors as introduced by Bernardin et al. (2009). Although this approach may require significantly more effort in terms of data collection and processing, potentially diminishing the appealing trade-off of minimal effort for log-likelihood gains, it could offer enhanced benefits that extend to subsequent choice models. Combining this advanced formulation with previous suggestions may provide a deeper understanding of how such enhancements affect the entire transport modeling framework, ultimately leading to more accurate and robust models for policy analysis and decision making.

In the long term, to improve predictive performance, research should focus on ensuring that models accurately represent the decision-making systems they are trying to represent, rather than solely focusing on improving explanatory or predictive abilities. This can be achieved by developing and including theories to address fundamental flaws. Drawing from relevant fields, such as psychology or other behavioral science, to further refine models, making them more representative of actual decision-making processes. Such an approach would help use simple models and data more efficiently to build robust models. This approach may even provide insight into how to make complex models more computationally feasible and data-efficient. By prioritizing behavioral representativeness, models naturally achieve better explanatory and predictive performance.

.

Bibliography

Ben-Akiva, M. (1974). Structure of passenger travel demand models.

Bernardin, V., Chen, J., and daniels, c. (2018). How-to: Model destination choice. Technical report, United States Department of Transportation Federal Highway Administration.

Bernardin, V. L., Koppelman, F., and Boyce, D. (2009). Enhanced destination choice models incorporating agglomeration related to trip chaining while controlling for spatial competition. *Transportation Research Record: Journal of the Transportation Research Board*, 2132:143–151.

Bhat, C., Govindarajan, A., and Pulugurta, V. (1998). Disaggregate attraction-end choice modeling formulation and empirical analysis. *Transportation Research Record*.

Bradley, M. A., Metro, P., Bowman, J. L., and Systematics, C. (1998). A system of activity-based models for portland, oregon. Technical report, U.S. Department of Transportation and Environmental Protection Agency. USDOT Report DOT-T-99-02. Produced for the Travel Model Improvement Program of the U.S. Department of Transportation and Environmental Protection Agency, Washington, D.C.

Bunch, D. S. (1987). Maximum likelihood estimation of probabilistic choice models. *SIAM Journal on Scientific and Statistical Computing*, 8.

Business.gov.nl (2024). Employment of young people: 15-year-olds.

Castiglione, J., Bradley, M., and Gliebe, J. (2014). *Activity-Based Travel Demand Models: A Primer*. Transportation Research Board.

Cats, O. (2022). Cieq6002 transport modelling and analysis lecture 2 slides: Trip generation and distribution.

Centraal Bureau voor de Statistiek (CBS) (2023). Gegevens per postcode. Accessed: October 29, 2024.

Centraal Bureau voor de Statistiek (CBS) (2024). Statline - dataset 83932ned. Accessed: October 29, 2024.

Clifton, K. J., Singleton, P. A., Muhs, C. D., and Schneider, R. J. (2016). Development of destination choice models for pedestrian travel. *Transportation Research Part A: Policy and Practice*, 94:255–265.

Coe, N. M., Kelly, P. F., and Yeung, H. W. C. (2007). *Economic Geography: A Contemporary Introduction*. Blackwell Publishing.

Daly, A. (1982). Estimating choice models containing attraction variables. *Transportation Research Part B: Methodological*, 16:5–15.

DANS (2024). Onderzoek onderweg in nederland. Accessed: 2024-06-02.

de Boer, E. and Blijie, B. (2006). Modelling school choice in primary education: An aid in school location planning - an exercise with the city of zwijndrecht schools. Report from Delft University of Technology, Department of Transport and Planning.

de Dios Ortúzar, J. and Willumsen, L. G. (2011). *Modelling Transport*. Wiley.

De Luca, S. and Cantarella, G. E. (2009). Validation and comparison of choice models. In Saleh, W. and Sammer, G., editors, *Travel Demand Management and Road User Pricing: Success, Failure and Feasibility*, pages 37–58. Ashgate Publications.

Dienst Uitvoering Onderwijs (2024). Duo open education data. Accessed: 2024-05-31.

Fotheringham, A. S. (1985). Spatial competition and agglomeration in urban modelling. *Environment and Planning A: Economy and Space*, 17.

Fox, J., Daly, A., Hess, S., and Miller, E. (2014). Temporal transferability of models of mode-destination choice for the greater toronto and hamilton area. *Journal of Transport and Land Use*, 7:41–62.

Fox, J. and Hess, S. (2010). Review of evidence for temporal transferability of mode-destination models. *Transportation Research Record*.

Guevara, C. A., Chorus, C. G., and Ben-Akiva, M. E. (2016). Sampling of alternatives in random regret minimization models. *Transportation Science*, 50:306–321.

Hauser, J. R. (1978). Testing the accuracy, usefulness, and significance of probabilistic choice models: An information-theoretic approach. *Operations Research*, 26.

Ho, C. Q. and Hensher, D. A. (2016). A workplace choice model accounting for spatial competition and agglomeration effects. *Journal of Transport Geography*, 51:193–203.

Jonnalagadda, N., Freedman, J., Davidson, W. A., and Hunt, J. D. (2001). Development of microsimulation activity-based model for san francisco: Destination and mode choice models. *Transportation Research Record: Journal of the Transportation Research Board*, 1777:25–35.

Kim, J. and Lee, S. (2017). Comparative analysis of traveler destination choice models by method of sampling alternatives. *Transportation Planning and Technology*, 40:465–478.

Koppelman, F. S. and Wilmot, C. G. (1982). Transferability analysis of disaggregate choice models. *Transportation Research Record*, 895.

Lekshmi, G. A., Landge, V., and Kumar, V. S. (2016). Activity based travel demand modeling of thiruvananthapuram urban area. *Transportation Research Procedia*, 17:498–505.

Luce, R. D. and Suppes, P. (1965). Preference, utility, and subjective probability. In Luce, R. D., Bush, R. R., and Galanter, E., editors, *Handbook of Mathematical Psychology*, volume 3, pages 249–410. John Wiley & Sons, New York.

Marschak, J. (1974). *Binary-Choice Constraints and Random Utility Indicators (1960)*, pages 218–239. Springer Netherlands.

Mauad, S. V. S. and Isler, C. A. (2024). Comparative analysis of discrete choice approaches for modeling destination choices of urban home-based trips to work. *Transportation Research Record: Journal of the Transportation Research Board*, 2678:1–16.

Metropoolregio Amsterdam (2023). About metropolitan region amsterdam. Accessed: 2024-05-26.

Nuffic (2024). Education in the netherlands. Accessed: 2024-06-07.

Object Vision (2024). Objectvision github repository. Accessed: 2024-11-27.

ObjectVision (2023). Openstreetmaps. Accessed: 2024-09-08.

ObjectVision (2023). Pc4 travel time matrix for nederland van boven. Accessed: September 8, 2024.

Onderzoek en Statistiek Amsterdam (2024). About metropolitan region amsterdam. Accessed: 2024-05-31.

Organisation for Economic Co-operation and Development (OECD) (2023). Data explorer - oecd wise database: Water indicators. Accessed: October 29, 2024.

Parady, G., Ory, D., and Walker, J. (2021). The overreliance on statistical goodness-of-fit and under-reliance on model validation in discrete choice models: A review of validation practices in the transportation academic literature. *Journal of Choice Modelling*, 38:100257.

Parody, T. E. (1977). Analysis of predictive qualities of disaggregate modal-choice models. *Transportation Research Record*.

Sanko, N. and Morikawa, T. (2010). Temporal transferability of updated alternative-specific constants in disaggregate mode choice models. *Transportation*, 37:203–219.

Shiftan, Y. (1998). Practical approach to model trip chaining. *Transportation Research Record: Journal of the Transportation Research Board*, 1645:17–23.

Statistics Netherlands (CBS) (2019). Onderweg in nederland (odin) 2018: Onderzoeksbeschrijving. Retrieved from Onderweg in Nederland (ODiN) 2018 report.

Statistics Netherlands (CBS) (2024). Dutch national travel survey. Accessed: November 6, 2024.

Suppes, P. and Luce, R. D. (1961). Individual choice behavior a theoretical analysis. *Journal of the American Statistical Association*, 56:172.

Sá, C., Florax, R. J. G. M., and Rietveld, P. (2004). Determinants of the regional demand for higher education in the netherlands: A gravity model approach. *Regional Studies*, 38:375–392.

Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46:234–240.

Train, K. (1978). A validation test of a disaggregate mode choice model. *Transportation Research*, 12.

Transportation Forecasting Resource (2024). Destination choice models. `https://tfresource.org/topics/Destination_Choice_Models.html`. Accessed: 10th March 2024.

van Welie, L., Hartog, J., and Cornelisz, I. (2013). Free school choice and the educational achievement gap. *Journal of School Choice*, 7:260–291.

A

ABM Framework

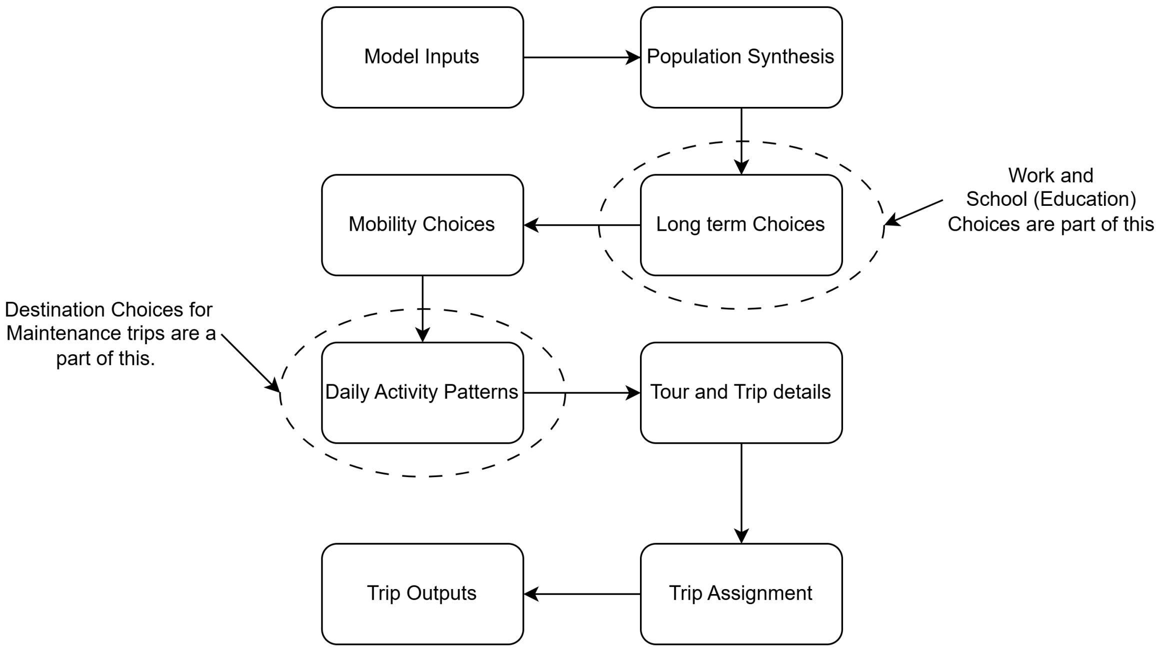


Figure A.1: Activity Based Modelling Framework (Castiglione et al., 2014)

B
Data Structure

B.0.1. Relevant ODiN variables

Table B.1: Numerical and Binary Variables

ODiN Variable	Meaning	Value Range	Category	Variable Type
leeftijd	Age	1-99	Person	Numerical
hhpers	No. of people in household	1-9 10: $\geq$ 10 11: Unknown	Household	Numerical
oprijbewijs	Car Driving License Availability	0: No 1: Yes 2: Unknown	Person	Binary

Table B.2: Nominal Variables

ODiN Variable	**Meaning**	**Value**	**Category**	**Note**
geslacht	Gender	1: Male 2: Female	Person	Could be converted into a binary variable for 0: Male 1: Female
maatspart	Social Participation	1: Employed 12-30 hrs/week 2: Employed 30+ hrs/week 3: Own Household 4: Scholar/student 5: Unemployed	Person	Could be used instead of Paid work or vice versa;
herkomst	Migration Background	1: Dutch 2: Western 3: Non-western 4: Unknown	Person	
ovstkaart	Possession of Student OV Chipcard	0: No 1: Weekly Pass 2: Weekend Pass 3: Unknown 4: NA/OP 15 years or 40 years	Person	
hhsam	Household Composition	1:Single 2:Pair/Couple 3:Couple+Children 4:Couple+Children+Others 5:Pair+Others 6:Single parent+Children 7:Single parent+Children+ Others 8:Other Household	Household	Useful for Maintenance trips

Table B.3: Ordinal Variables

ODiN Variable	**Meaning**	**Value**	**Category**	**Note**

Continued on next page

Table B.3: Ordinal Variables (Continued)

opbezitvm	Owned Transport Means	0: No Means of transport 1: ≥3 Cars 2: ≥ 2 Cars 3: ≥1 Car 4: ≥ 1 Motored Vehicle 5: ≥ 1 Moped 6: ≥ 1 Light Moped 7: Unknown	Person	This could be converted into a binary variable has_car where: if value in 1,2,3 has_car = 1 else 0 Alternatively, there is another variable called 'HHBezitVm' for Household transport ownership which could be used instead of this
betwerk	Paid Work	0: No gainful employment 1: 12 hrs/week 2: 12-30 hrs/week 3: ≥30 hours/week 4: Unknown 5: Not requested/OP below 15 years of age	Person	
hhgestinkg	Household Standardised Disposal Income level	1 First 10% Group 2: Second 10% 3-10 Groups of 10% 11 Unknown	Household	
fqnefiets	Bike Usage Frequency	1:Daily/Almost daily 2: Several times a week 3: Several times a month 4: Several time a year 5: Never/Rarely	Person	This could be converted into a binary variable has_bike where: if value in 1,2,3 has_bike = 1 else 0 Similarly there are variable for e-bikes and mopeds
sted	Urbanization level of Residence Municipality	1: Very High Urban 2: Strongly Urban	Person	

Continued on next page

Table B.3: Ordinal Variables (Continued)

		3: Moderately Urban 4: Less Urban 5: Non- Urban		
opleiding	Education Level	0: No training completed 1: Basic education 2: Lower vocational education 3: Secondary vocational education 4: Higher vocational education or University 5: Other Training 6: Unknown 7: Not requested/OP below 15 years of age	Person	

Table B.4: ODiN Variables for filtering trips

ODiN Variable	Meaning	Value Range	Category	Variable Type
kmotiefv	Travel Motive	1: To and From work 3: Services/Personal Care 4: Shopping 5: Education	Person	Nominal
doel	Purpose	1: Going home 2: Work 7:Educational course		

B.0.2. Employment data (Onderzoek en Statistiek Amsterdam, 2024)

Table B.5: Employment data for the MRA region

Column Name	Sector	gebiedsniveau	gebiedscode	gebiedsnaam	jaar	aantal
Meaning	Sector	Area level	Area code	Area name	Year	Number
Values	• Total • Wholesale and Retail • Horeca • Financial Institutions • Utilities • Government • Health and welfare • Other services	Gemeenten	Gemeente Code Corresponds to the codes used in ODiN data. The number of jobs is also available at the PC4 level.	Name of Muncipality	2018-2023	No. of Jobs[10]

[10]The number of jobs refers to the total number of full-time, part-time, and, temporary workers

B.0.3. Education enrollment data (Dienst Uitvoering Onderwijs, 2024)

Table B.6: Enrollment data for Primary education

Column Name	Peildatum	Postcode vestiging	Gemeentenummer	Gemeentenaam	Soort_po	Totaal
Meaning	Reference data	Postal code of school	Municipality number.	Municipality name	type primary education	Total enrollments
Values	The moment to which the data relates. A reference date in 2013 refers to the 2013-2014 school year. Reference date in 2014 refers to the 2014-2015 school year, etc	PC6 codes	Municipality codes corresponding to the ones in ODiN data	Names of Municipality	The type of primary education: • bo: primary education • sbo: special primary • vso: secondary special • so: special education	Total number of students enrolled at the school

Table B.7: Enrollment data for Secondary education

Column Name	Instellingsnaam Vestiging	Gemeentenummer	Gemeentenaam	Totaal Aantal Leerling
Meaning	Institution Name	Municipality number.	Municipality name	Total enrollments
Values	Name of the educational institute. It can be used to get the PC4 codes	Municipality codes corresponding to the ones in ODiN data	Names of Municipality	Students enrolled for VMBO, HAVO, VWO with VAVO counts as separate at the institute

Table B.8: Enrollment data for Vocational and Higher education

Vocational Education				
Column Name	Instellingsnaam Vestiging	Gemeentenummer	Gemeentenaam	Aantal Leerling
Meaning	Institution Name	Municipality number.	Municipality name	Enrollments
Values	Name of the educational institute. It can be used to get the PC4 codes	Municipality codes corresponding to the ones in ODiN data	Names of Municipality	Students enrolled for various types of MBOs. It can be summed to calculate total enrollments. Data available from 2019 to 2023
Higher education				
	Same as above	Same as above	Same as above	Students enrolled for HBOs and WOs distinguished by course and gender at the institute. It can be summed to calculate total enrollments Data available from 2019 to 2023

C

GeoDMS Input Parameters

Parameter	Value	Units	Description
OngelijkvloersPenalty	2[min_f]	min_f	Time penalty when changing at a stop with a lot of vertical distance
WalkingSpeed_kmhr	4.5[km_hr]	km_hr	
BikingSpeed_kmhr	14[km_hr]	km_hr	
WalkingSpeed	WalkingSpeed_kmhr / 3600[s_f / hr_f] * 1000[m / km]	m_s	
BikingSpeed	BikingSpeed_kmhr / 3600[s_f / hr_f] * 1000[m / km]	m_s	
MaxCarSpeed	130[km_hr]	km_hr	
MaxCarSpeed_limit	100[km_hr]	km_hr	
PedestrianDefaultSpeed	WalkingSpeed_kmhr	km_hr	
BikeDefaultSpeed	BikingSpeed_kmhr	km_hr	
CarDefaultSpeed	50[km_hr]	km_hr	
CarDefaultSpeed_low	30[km_hr]	km_hr	
Ferry_Speed	25[km_hr]	km_hr	
TransferEffectiveSpeed	value(4[km_hr] / 1.2f, m_s)	m_s	The transfer walking speed: X km/hour / 1.2 (correction for Manhattan distances) and then converted to meter/sec

D

Metropolitan Region of Amsterdam

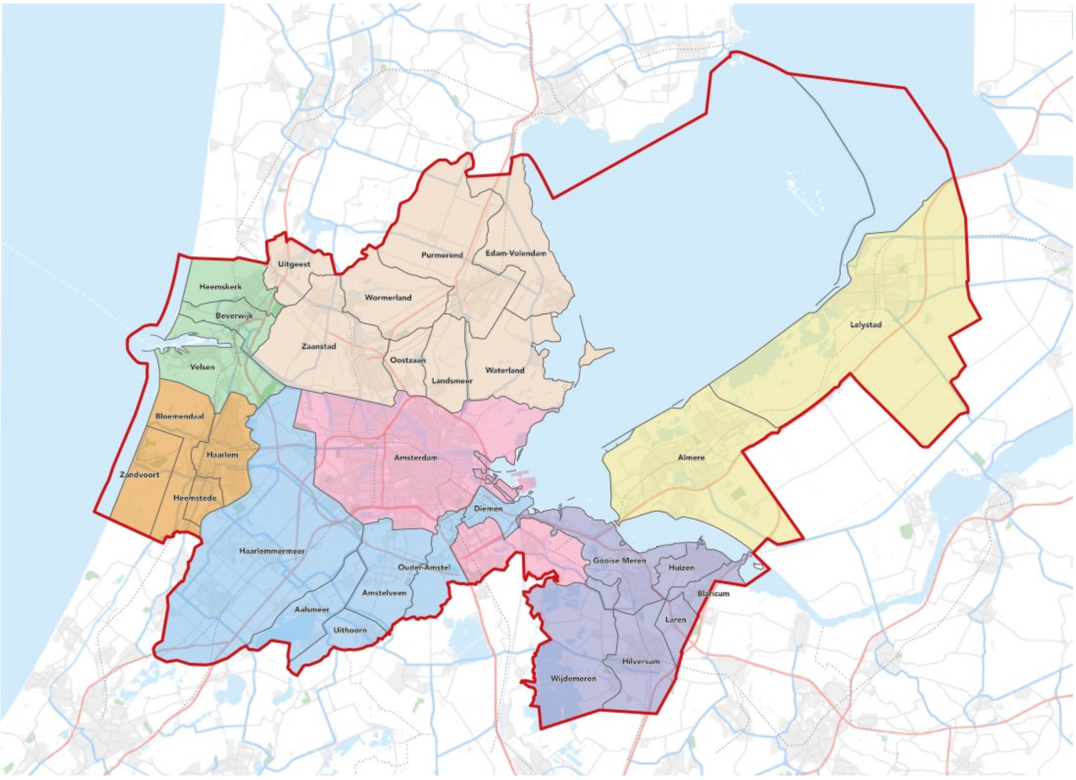


Figure D.1: Metropolitan region of Amsterdam: 30 Municipalities across 7 sub-regions
(Metropoolregio Amsterdam, 2023)

The MRA is subdivided into 7 sub-regions. The 30 municipalities along with their Gemeente Number in each of the sub-regions are as follows:

South Kennemerland
- 392: Haarlem
- 397: Heemstede
- 377: Bloemendaal
- 473: Zandvoort

Ijmond
- 453: Velsen
- 372: Beverwijk
- 396: Heemskerk

Zaanstreek-Waterland
- 479: Zaanstad
- 439: Purmerend
- 385: Edam-Volendam
- 852: Waterland
- 880: Wormerland
- 450: Uitgeest
- 415: Landsmeer
- 431: Oostzaan

Amstelland-Meerlanden
- 394: Haarlemmermeer
- 362: Amstelveen
- 358: Aalsmeer
- 451: Uithoorn
- 384: Diemen
- 437: Ouder-Amstel

Almere-Lelystad
- 34: Almere
- 995: Lelystad

Gooi and Vechtstreek
- 402: Hilversum
- 1942: Gooise Meren
- 406: Huizen
- 1696: Wijdemeren
- 417: Laren
- 376: Blaricum

Amsterdam
- 363: Amsterdam

D.1. Zonal Size vs observed trip distributions

D.1.1. Total Job vs Work trip distributions

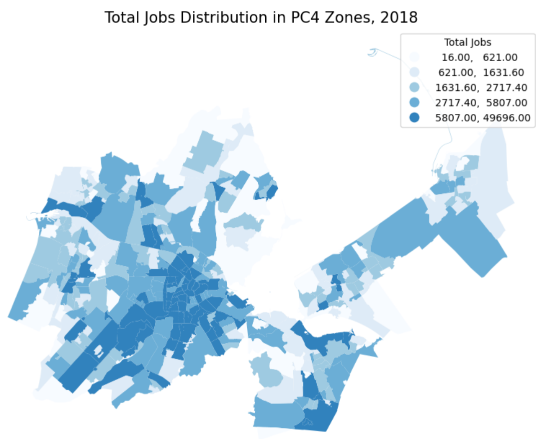


Figure D.2: Total Jobs Distribution (2018)

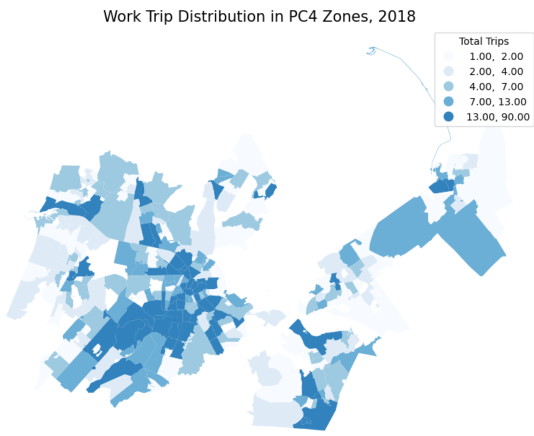


Figure D.3: Work Trip Distribution (2018)

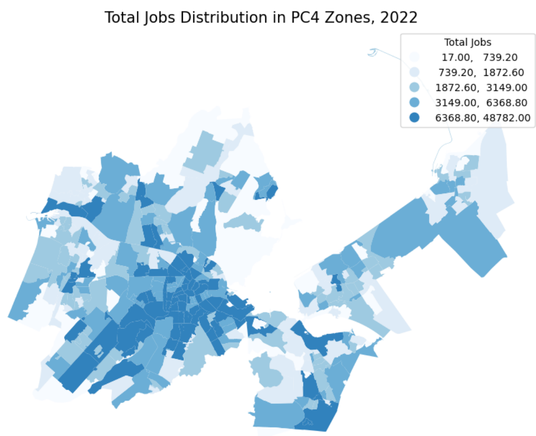


Figure D.4: Total Job Distribution (2022)

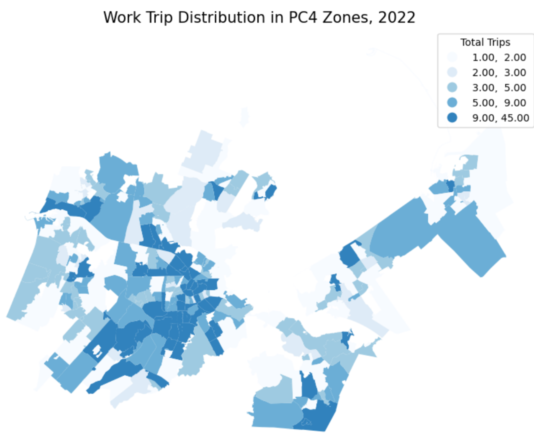


Figure D.5: Work Trip Distribution (2022)

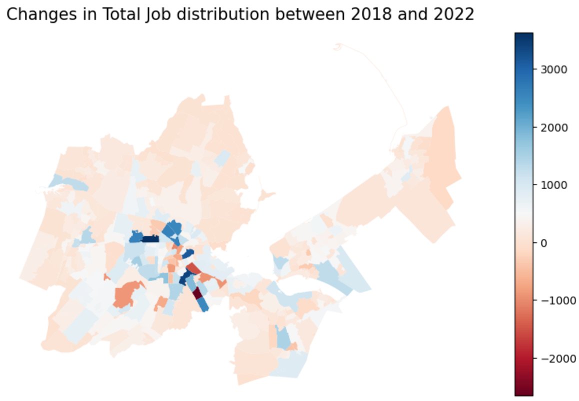


Figure D.6: Changes in Total Job distribution between 2018 and 2022

D.1.2. Retail and Service Job vs HBM trip distribution

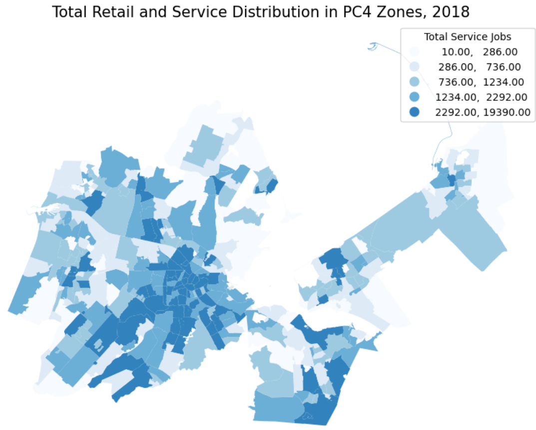


Figure D.7: Total Retail and Service Jobs Distribution (2018)

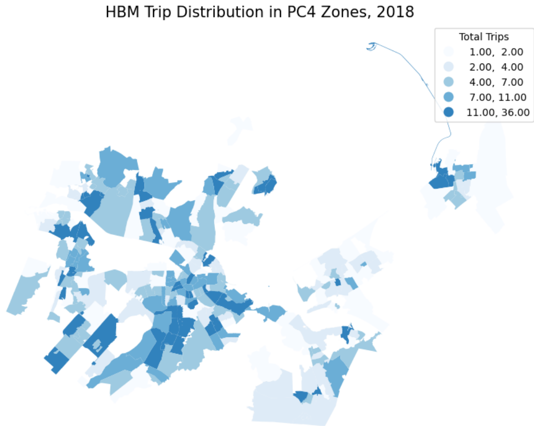


Figure D.8: HBM Trip Distribution (2018)

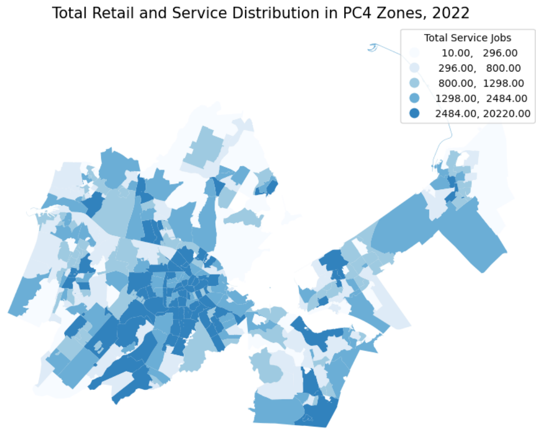


Figure D.9: Total Retail and Service Job Distribution (2022)

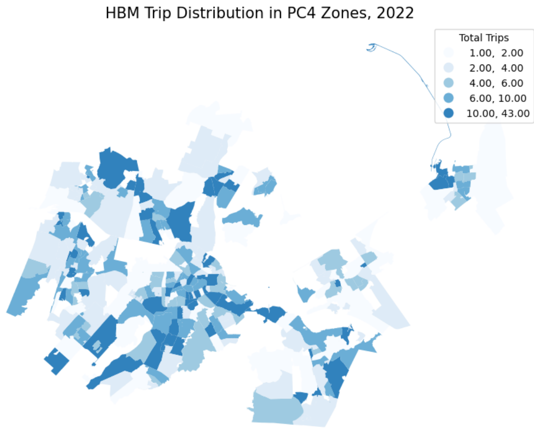


Figure D.10: HBM Trip Distribution (2022)

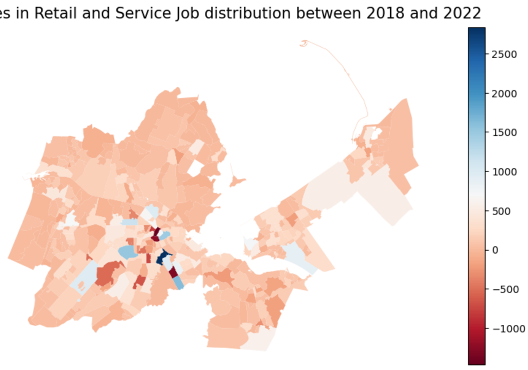


Figure D.11: Changes in Total Retail and Service Job distribution between 2018 and 2022

D.1.3. Secondary and above Enrollment vs trip Distributions

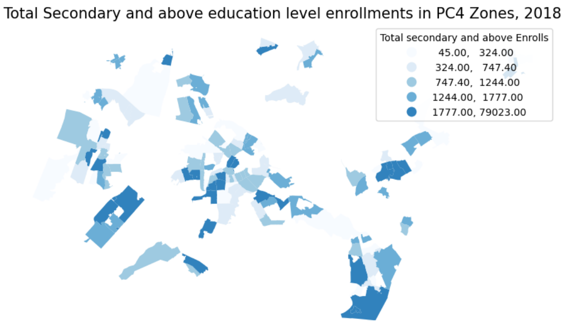


Figure D.12: Total Secondary and above Enrollment Distribution (2018)

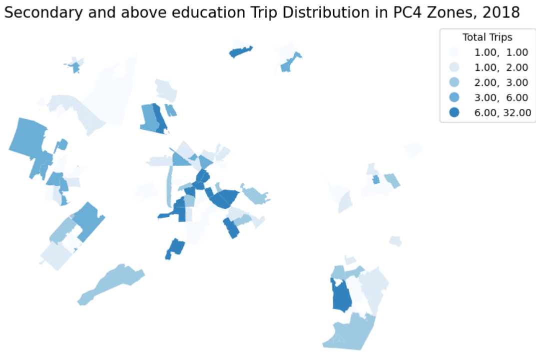


Figure D.13: Secondary and above Trip Distribution (2018)

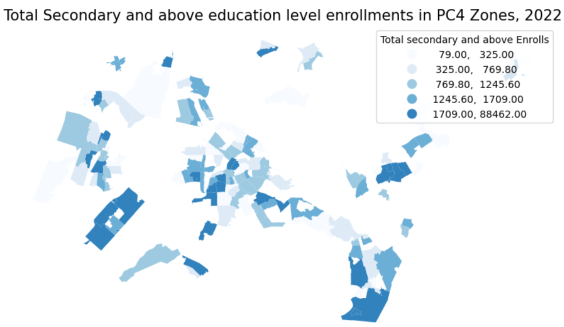


Figure D.14: Total Secondary and above Enrollment Distribution (2022)

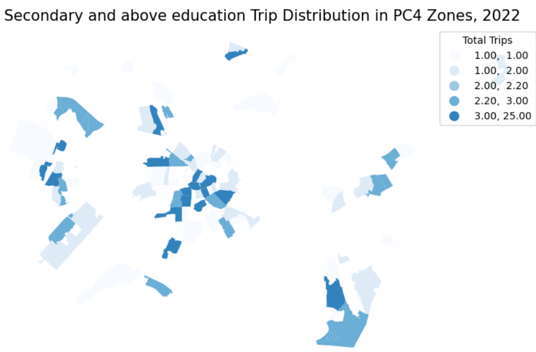


Figure D.15: Secondary and above Trip Distribution (2022)

D.1.4. Primary Enrollment vs trip distribution:

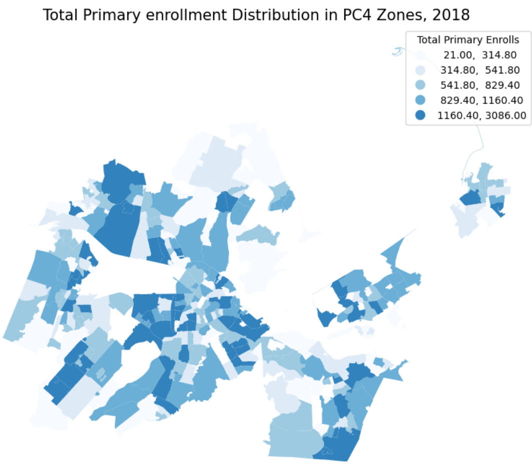


Figure D.16: Total Primary Education Enrollment Distribution (2018)

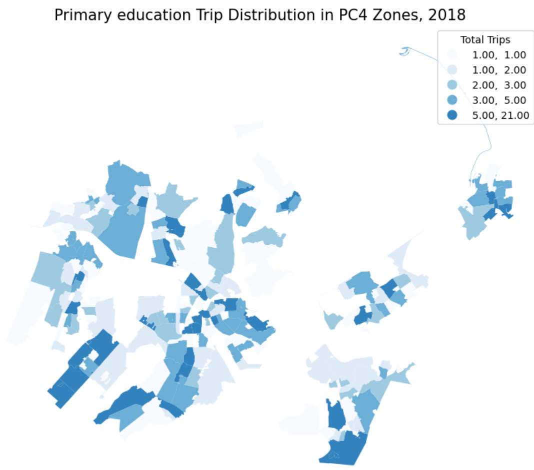


Figure D.17: Primary Education Trip Distribution (2018)

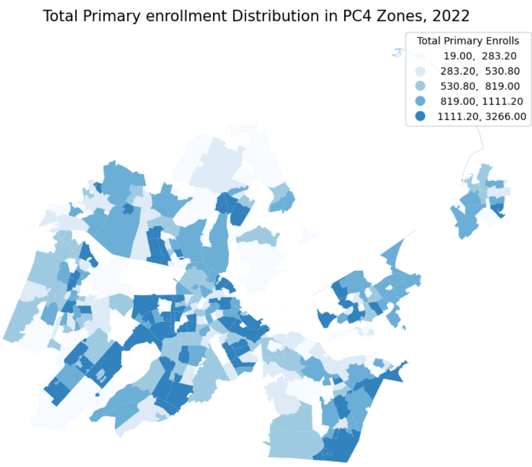


Figure D.18: Total Primary Education Enrollment Distribution (2022)

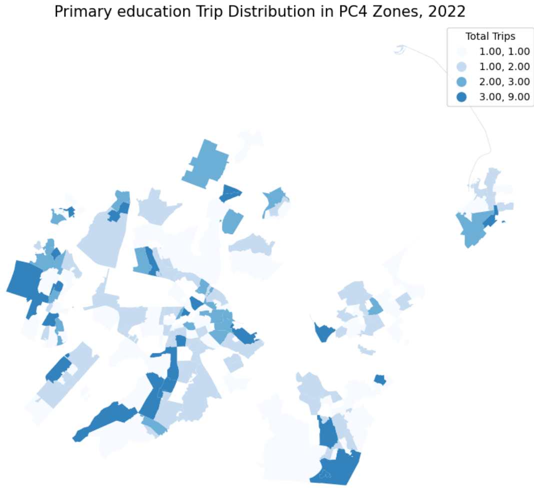


Figure D.19: Primary Education Trip Distribution (2022)

D.2. Changes in travel times between 2018 and 2022

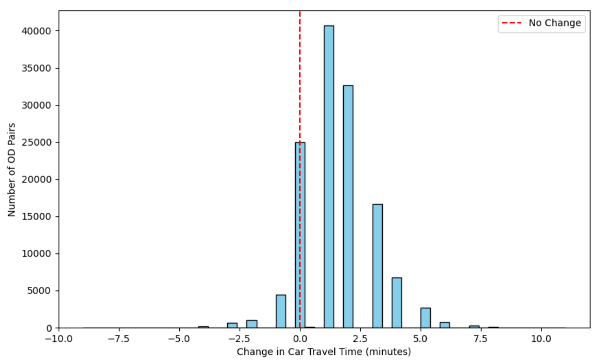


Figure D.20: Change in car travel time between PC4 zones in the MRA

D.3. % Intrazonal trips and average travel time per trip purpose

Table D.1: % Intrazonal and average travel times per trip purpose

Trip Purpose	2018		2022	
	% Intrazonal Trips	Average travel time (mins)	% Intrazonal Trips	Average travel time (mins)
Work	5%	21.83	6.36%	23.27
HBM	37.35%	7.80	36.52%	7.81
Secondary Education and above	6.59%	17.18	5.43%	18.92
Primary Education	41.85%	16.37	45.82%	18.91

E

Dutch Education System

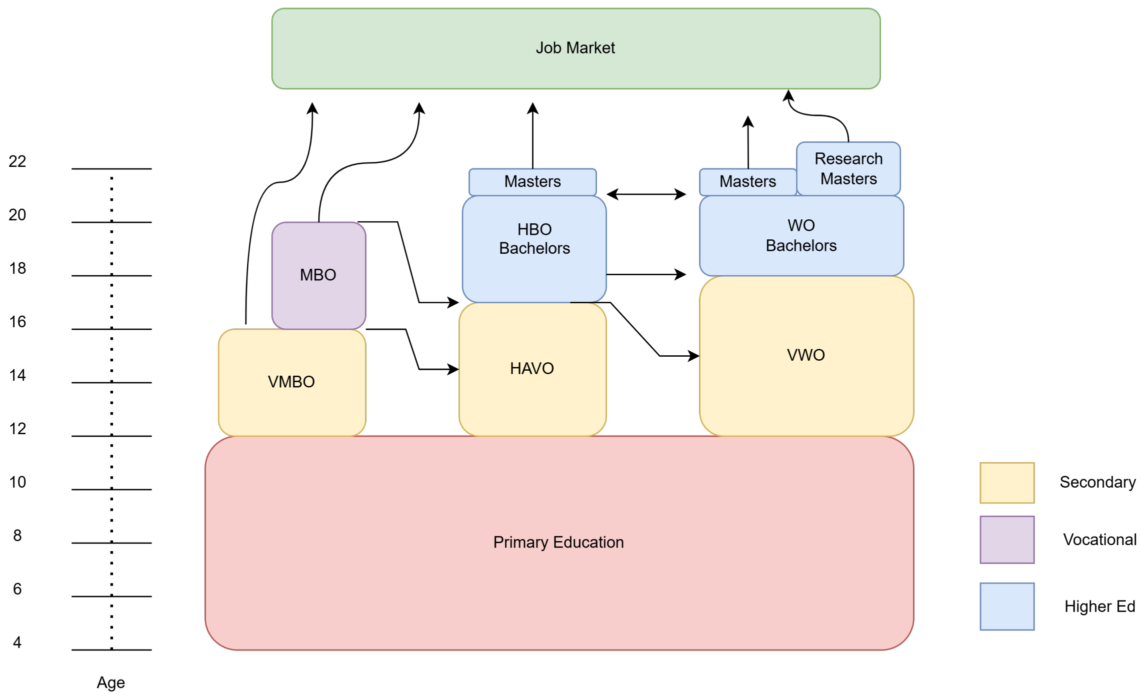


Figure E.1: Education system in the Netherlands (Nuffic, 2024)

F

Statistics per Trip purpose for Processed ODiN Data

F.1. Work

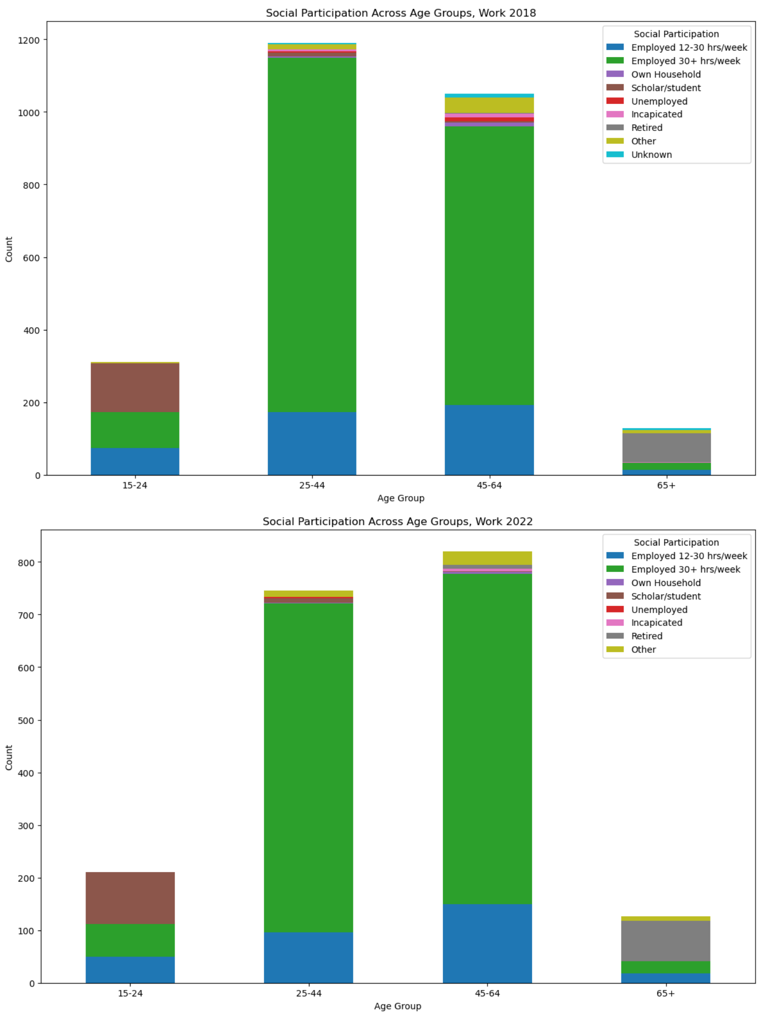


Figure F.1: Distribution of Social participation status across Age groups for work trips

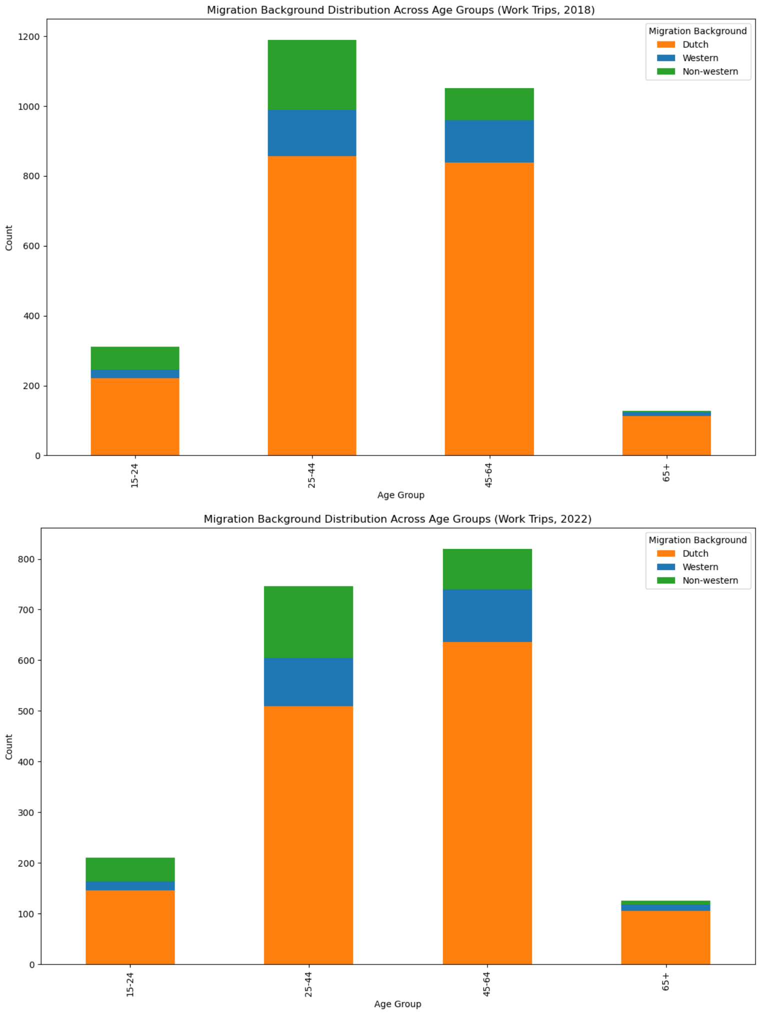
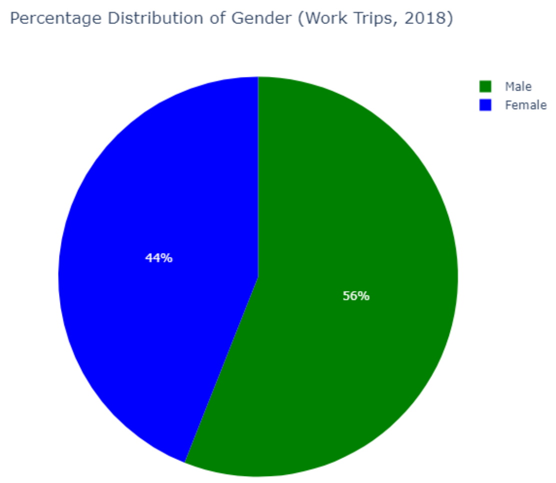


Figure F.2: Distribution of Migration background across Age groups for work purpose



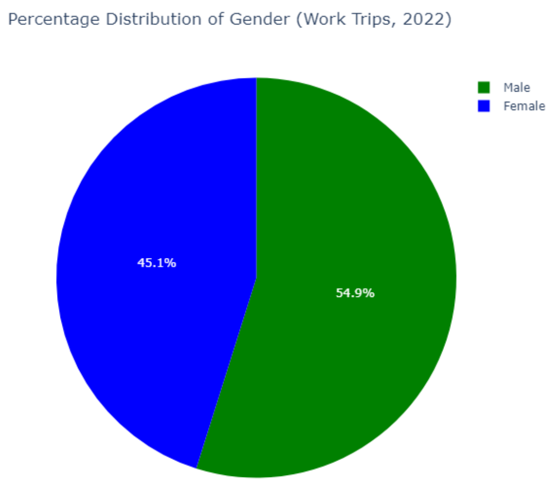


Figure F.3: Gender proportion amongst travelers for work purpose

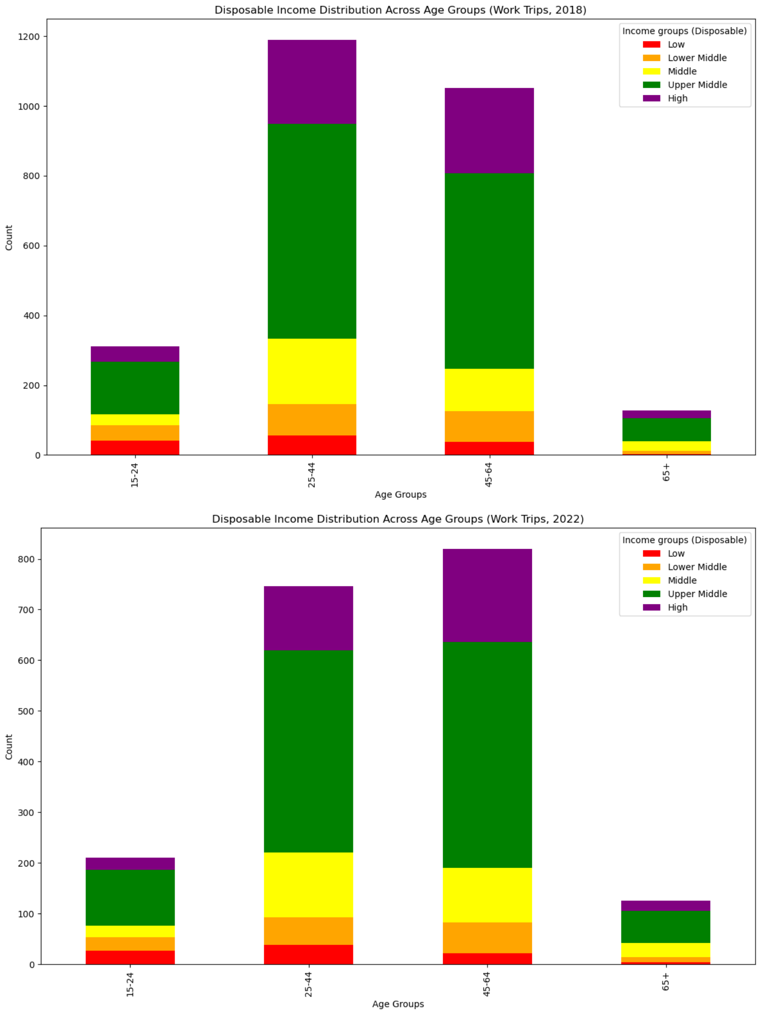


Figure F.4: Distribution of Disposable income level across Age groups (work)

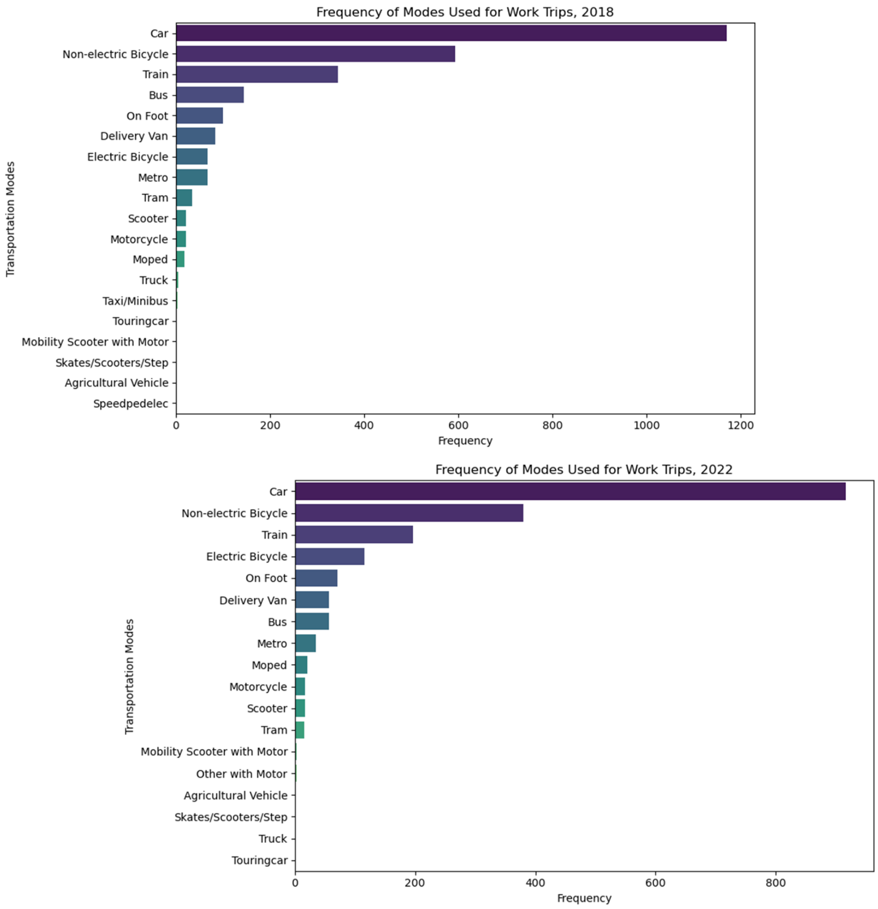
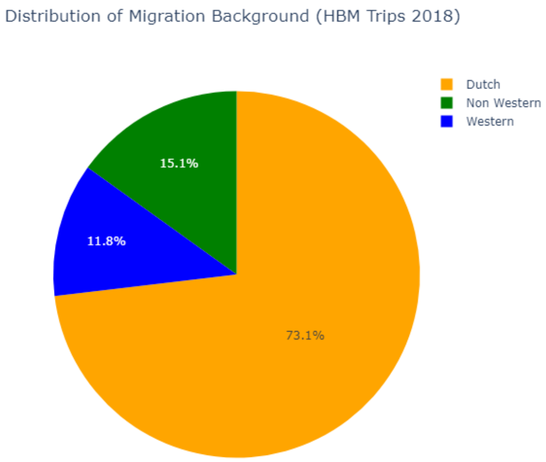


Figure F.5: Observed Mode Frequency Work trips 2018, 2022

F.2. HBM



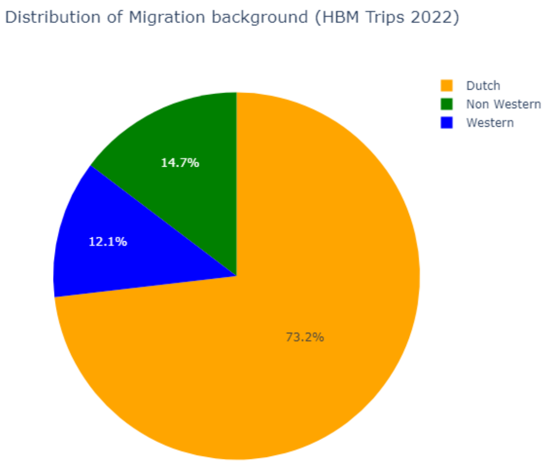


Figure F.6: Migration background proportion among travelers for HBM trips

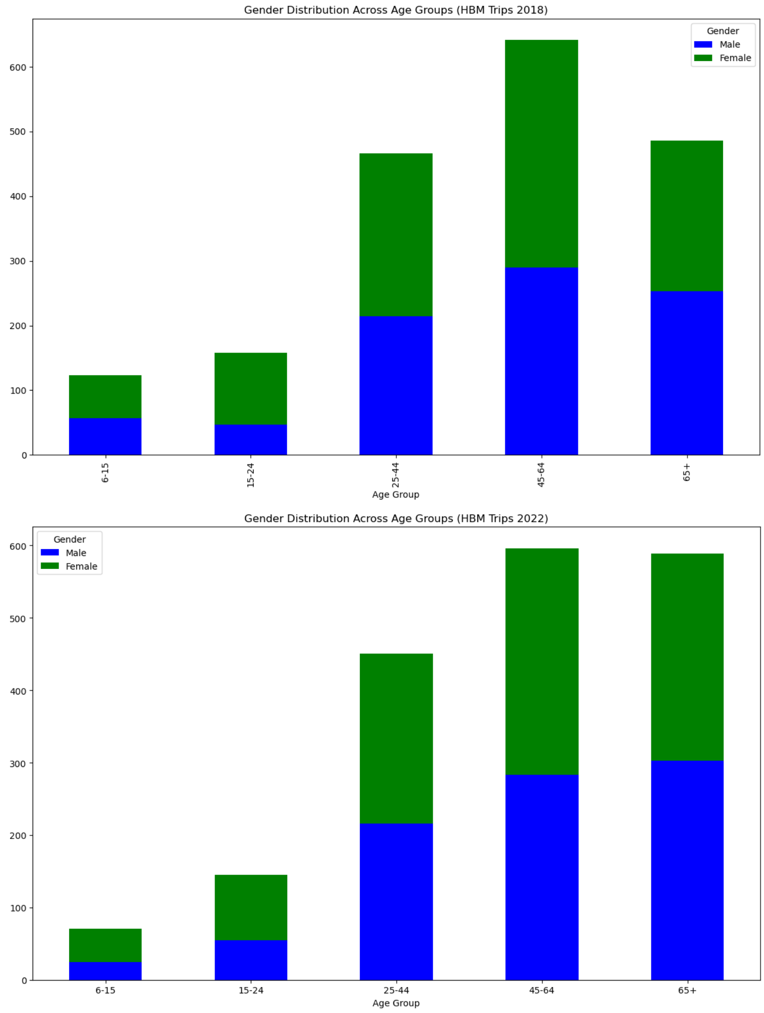


Figure F.7: Gender distribution across age groups for 2018 and 2022 (HBM trips)

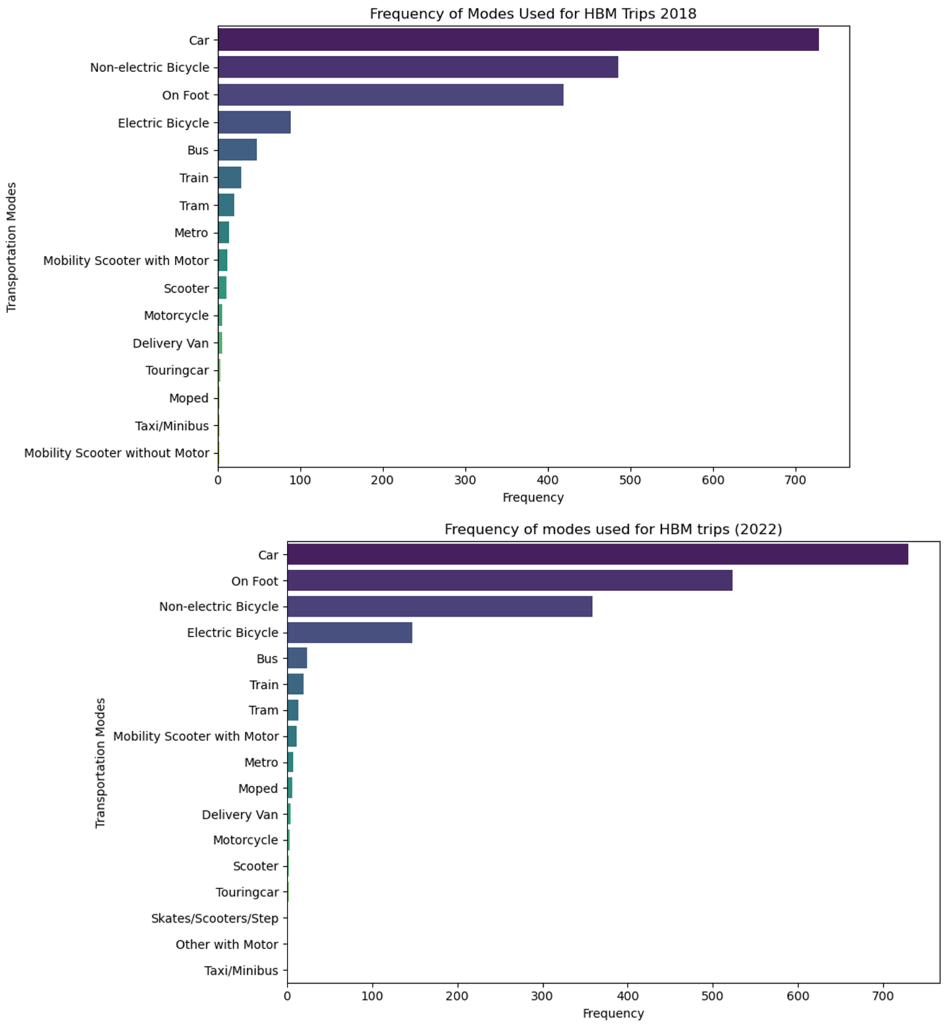


Figure F.8: Observed Mode Frequency HBM trips 2018, 2022

F.3. Primary education

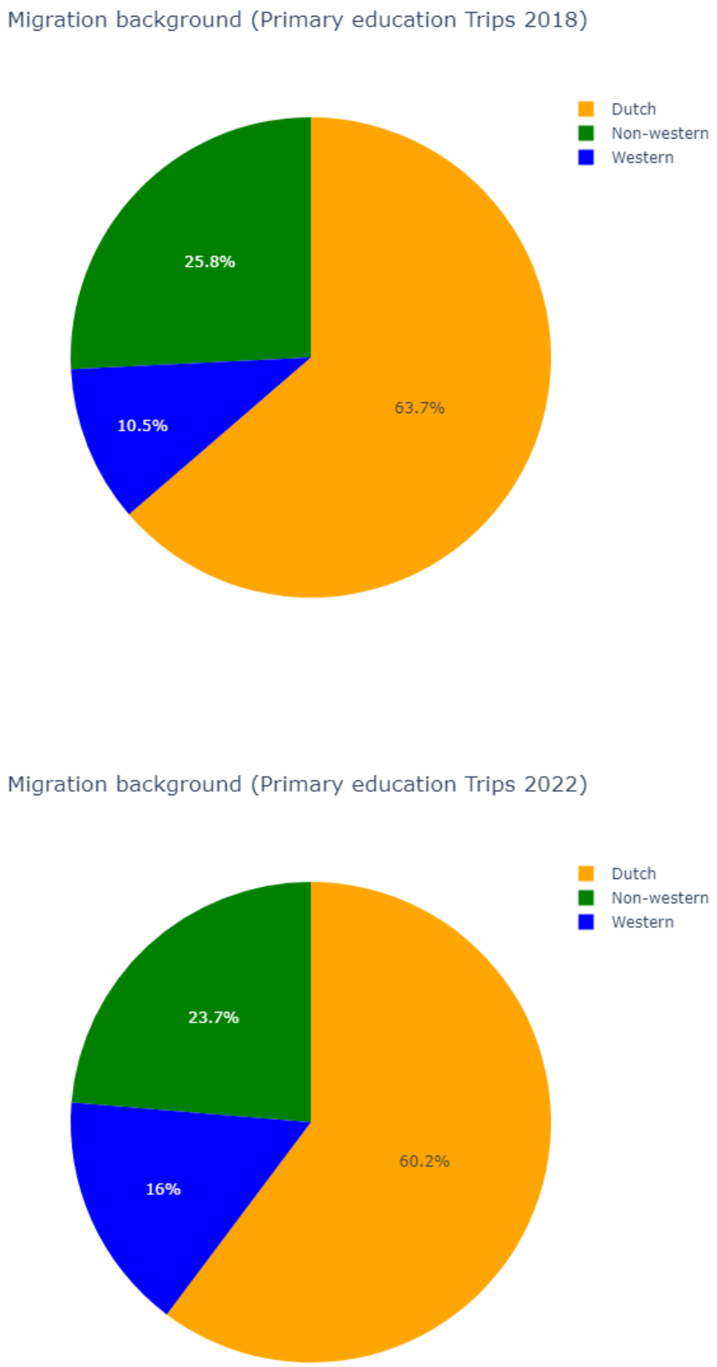


Figure F.9: Migration background distribution (Primary education trips) 2018, 2022

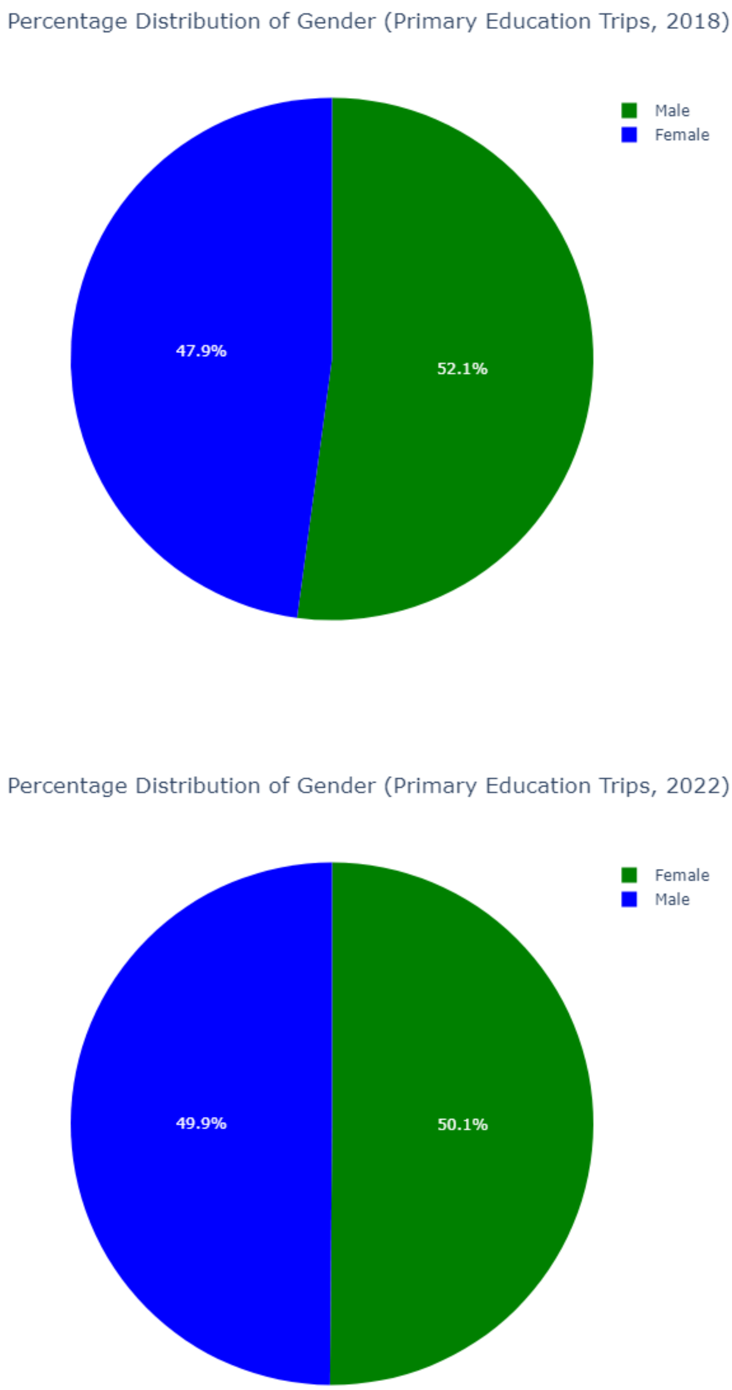


Figure F.10: Gender distribution (Primary education trips) 2018, 2022

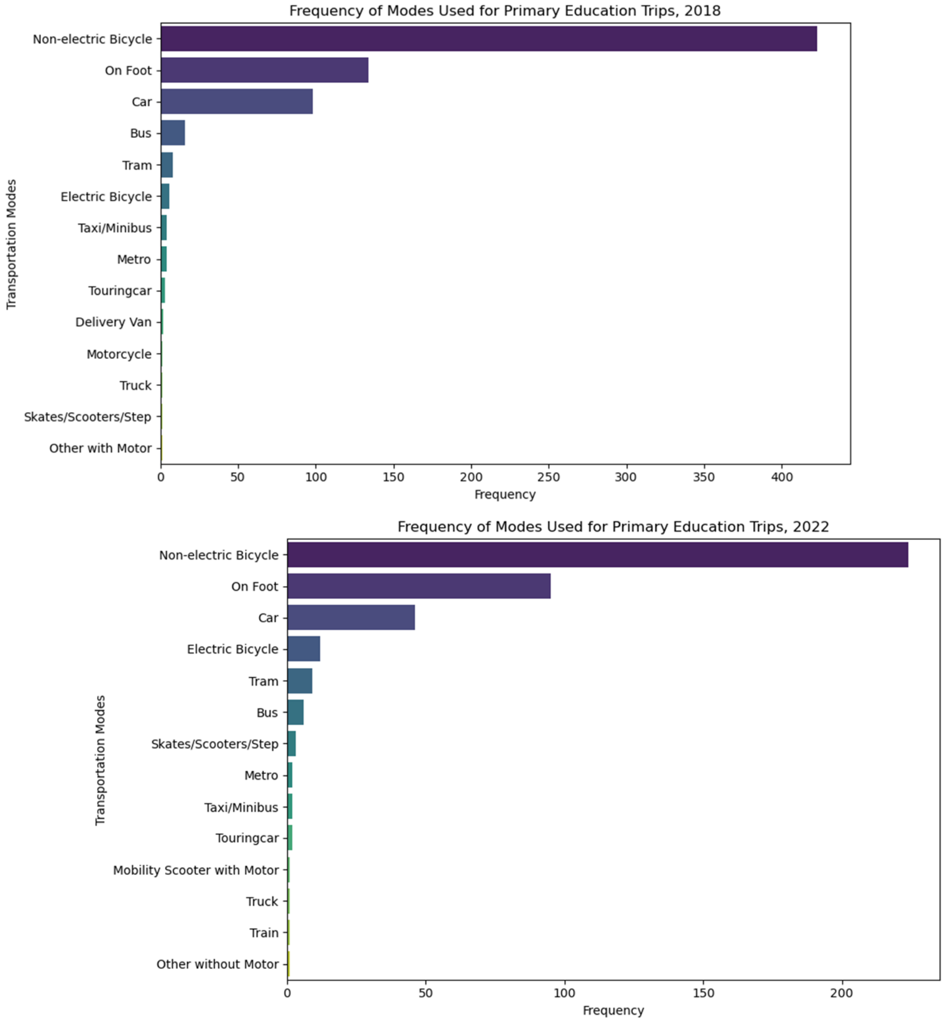


Figure F.11: Observed Mode Frequency Primary Education trips 2018, 2022

F.4. Secondary and above education trips

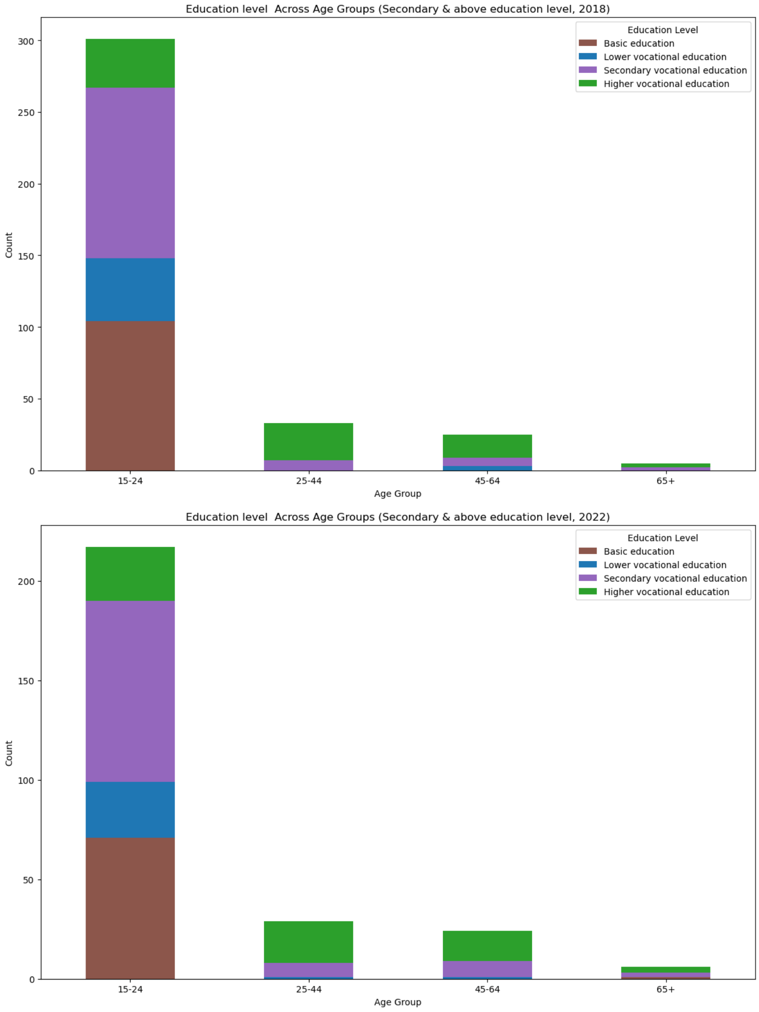


Figure F.12: Highest completed education distribution across age groups for 2018 and 2022 (Secondary and above trips)

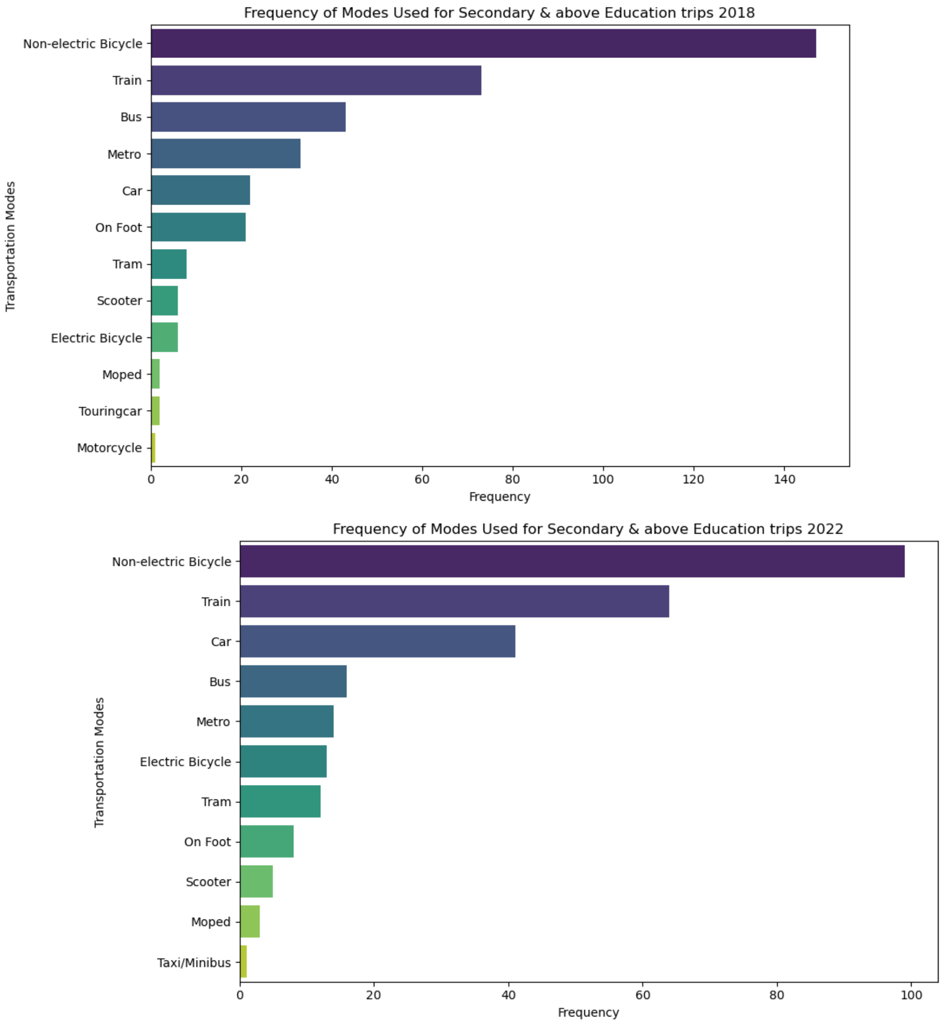


Figure F.13: Observed Mode Frequency Secondary Education trips 2018, 2022

G

Scientific paper

Impact of Spatial Competition and Agglomeration Effects on the Temporal Transferability of Destination Choice Models

A Case Study on the Metropolitan Region of Amsterdam

Rohan Menezes

Abstract

To ensure transport models are effective tools for planning, they must not only adequately explain current travel choices but also maintain predictive accuracy over forecast horizons while being computationally feasible for practical use. One simple approach to achieving this is by including behavioral theories in transport models to improve the behavioral representation of models. One such theory is Spatial Competition and Agglomeration Effects (SC&AE). This theory examines how opportunities present nearby influence the attractiveness of a destination. This influence can be either positive (Agglomeration) or negative (Spatial Competition). Although SC&AE is widely recognized in the literature for enhancing the explanatory power of computationally simple Multinomial Logit (MNL) destination choice models, its impact on the temporal transferability of these models remains unexplored.

This study assesses the impact of SC&AE on MNL destination choice models for home-based maintenance, work, and education trips in the Metropolitan Region of Amsterdam on a 5-year short-term forecast horizon (2018-2022) using Dutch National Travel Survey Data (ODiN). The findings indicate that SC&AE has a positive but limited effect on the temporal transferability of these models. This impact decreases with decreasing traveler autonomy and ease of switching to alternate destinations. In terms of percentage improvement, Transfer index metric shows an exaggerated impact (35.2% - 286.36%) while the other performance indicators, such as Fitting Factor (1.1% - 3.6%) and % Correct predictions (0.14%-2.05%) confirm that the impact is limited.

Keywords: Destination Choice models, Spatial Competition, Agglomeration, Temporal Transferability

1. Introduction

Transport planning relies heavily on models for forecasting travel behavior, owing to the long-term impact of policies and their resource-extensive execution. To ensure transport models are effective tools for planning, they must not only adequately explain current travel choices but also maintain predictive accuracy over forecast horizons while being computationally feasible for practical use. A simple approach to achieving this is by including behavioral theories in transport models to improve the behavioral representation of models.

Although these theories may enhance the model's ability to explain current travel choices, they do not necessarily improve their forecasting abilities due to the risk of overfitting. Overfitted models tend to explain random noise rather than the signal in the data (Parady et al., 2021). Thus, overfitting negatively impacts the model's temporal transferability, i.e., its ability to maintain predictive accuracy across forecast horizons.

One such theory is Spatial Competition and Agglomeration Effects (SC&AE). This theory examines how opportunities present nearby influence the attractiveness of a destination. This influence can be positive (Agglomeration) or negative (Spatial Competition). Accounting for this spatial heterogeneity remedies the popular Independence of Irrelevant Alternatives (IIA) assumption of Multinomial Logit (MNL) destination choice models (DCM). Relaxing the IIA assumption improves the model's explanatory power for destination choices. Specifically, this improvement occurs when SC&AE are accounted for through accessibility measures in the utility specification and is well noted in transport literature across various trip purposes and geographical regions over the years (Bernardin et al. (2009); Ho and Hensher (2016); Sá et al. (2004)).

Thus, given its long-term applicability in various contexts, these effects are potentially essential in explaining travelers' destination choices and, therefore, can contribute significantly to MNL DCM's temporal transferability. Yet, its impact on the temporal transferability of these models remains unexplored.

This research seeks to provide quantitative evidence of the SC&AE impact on the temporal transferability of MNL DCMs. Thus, it assesses the validity of SC&AE in explaining destination choices to justify its inclusion in MNL DCMs. Such a step towards a more holistic validation enables assessing whether SC&AE captures travel behavior that drives destination choices rather than merely capturing behavior contextually in the travel data.

2. Literature Overview

2.1. Multinomial Logit Models for Destination Choices

Trip Distribution models estimate the flow of trips between origin and destination (OD) pairs as a function of travel impedance and size. They can broadly classified into aggregate and disaggregate models.

Aggregate models distribute trips based on observed patterns for groups of travelers or average relationships at the zonal level. A common aggregate method is the gravity model. Drawing an analogy to Newton's law of gravitation, it estimates the trip flows proportional to the product of origin and destination attractiveness and inversely proportional to travel impedance (de Dios Ortúzar and Willumsen, 2011).

Disaggregate models, on the other hand, are based on observed choices at the individual traveler or household level. The discrete choice model based on the utility maximization theory is the most commonly used type of disaggregate model. Because these disaggregated models are based on theories of individual behavior and do not rely on physical analogies, they have the potential advantage of modeling behavior more realistically and are more likely to be robust in explaining behavior in time and space. Among the many types of discrete choice models, the multinomial logit (MNL) model is computationally the simplest and most practical (de Dios Ortúzar and Willumsen, 2011).

However, previous research highlights a key limitation of MNL and gravity models: their popular unrealistic Independence of Irrelevant Alternatives (IIA), which assumes uniform, equal competition among all destination alternatives. Because this assumption does not hold for correlated alternatives, which is usually the case for nearby destinations (Tobler, 1970), it results in unrealistic predictions, as nearby destinations are more likely to be similar and compete more strongly with each other than distant ones. To address this limitation, one solution proposed by previous researchers is to enable MNL models to account for Spatial Competition and Agglomeration Effects (SC&AE) (Bernardin et al., 2009).

2.2. Spatial Competition and Agglomeration Effects in Destination Choices

Spatial Competition and Agglomeration Effects (SC&AE) consider how the spatial distribution of opportunities across destination alternatives affects an individual traveller's destination choice. Spatial competition arises when opportunities in nearby zones decrease the attractiveness of a destination to a traveler. On the other hand, if the opportunities increase the destination's attractiveness to the traveler, it is called the agglomeration effect.

One of the approaches introduced by Fotheringham (1985) of including SC&AE in MNL models is to use a Hansen-type accessibility index to include information about other destination alternatives. This accessibility index for destination zone j (A_j) captures the net effect of SC&AE, using a attraction size variable (R_z), such as employment in other destinations z and travel impedance between the zone j and other destinations z in the study area:

$$A_j = \ln \sum_{z \neq j} \frac{R_z}{c_{jz}} \tag{1}$$

The utility function for destination j for a traveler with origin zone i can then be extended and further specified as a linear function of A_j and its parameter estimate β_A to account for SC&AE, as follows:

$$P_{ij} = \frac{e^{v_{ij} + \beta_A A_j}}{\sum_{j'} e^{V_{ij'} + \beta_A A_{j'}}} \tag{2}$$

If $\beta_A < 0$, zones close to other opportunities have lower utility, indicating that competition effects dominate. If $\beta_A > 0$, zones close to other opportunities have a higher utility, indicating that agglomeration effects dominate. If $\beta_A = 0$, then there are no SC&AE or equally strong agglomeration and competition effects that cancel each other. Using this accessibility measure, which includes information about alternative destinations, the MNL model's IIA assumption does not apply. (Ho and Hensher, 2016).

SC&AE in DCMs has been a major focus of travel behavior research. Various studies have explored these effects in diverse contexts such as educational choices, workplace locations, and maintenance trip destinations in various regions globally. Building on the work of Fotheringham (1985), Bhat et al. (1998) included SC&AE as a single accessibility variable capturing the net effect to explain destination choices for home-based shopping trips in the Boston Metropolitan area (BMA). They found a highly significant negative parameter, indicating dominant spatial competition effects. The study also explored sociodemographic interactions with travel impedance and found that older adults and women were more sensitive to travel impedance for work trips. At the same time, higher-income travelers were more willing to travel longer distances for work. However, using a single accessibility index reveals net agglomeration and competition effects but cannot identify the source of these effects or the presence of non-dominating effects. To address this, Bernardin et al. (2009) introduced Agglomeration and Competing Destination Choice (ACDC) models, separating the effects using accessibility measures for complements and substitutes to analyze home-based maintenance (HBM) and home-based other (HBO) trips.

The study compares the results with the single accessible index estimated parameter and finds that numerically, they highlight a dominating spatial competition effect. In addition, the ACDC model finds that spatial competition for HBM and HBO trips arises from substitutes (employment within similar sectors) and agglomeration from complements (employment across different sectors) Ho and Hensher (2016) adapted the ACDC model to study workplace location choices (WLC) in the Sydney Greater Metropolitan Area (SGMA). The findings showed that for work trips, competition effects were driven by complements, unlike the substitutes identified in non-work trips by Bernardin et al. (2009). Similar to Bernardin et al. (2009), Ho and Hensher (2016) finds that the results from the ACDC and a single accessible index are numerically the same.

For educational choices, Sá et al. (2004) included SC&AE as a single accessibility index, naming it as "Centrality Index" within their production-constrained gravity model for trip distribution, to explain university location choices among high school graduates. The findings indicate that universities in densely populated areas experience competition effects, with a negative centrality index suggesting that nearby institutions compete for students, rather than benefit from agglomeration.

2.3. Factors affecting Temporal Transferability of Choice models

In travel behavior, transferability can be defined as " The ability of a model developed in one context to explain behavior in another, assuming the underlying theory is equally applicable in both contexts" (Fox et al., 2014). Thus, temporal transferability is the model's ability to maintain its accuracy and reliability over a forecasting horizon without requiring extensive recalibration. Assessing temporal transferability helps determine how well a model can adapt to changes over time, providing confidence in its predictions for transport planning and its ability as a tool to aid well-informed decision-making. This concept is particularly relevant in transportation planning, where models are used to predict future travel behaviors based on past data (Fox et al. (2014); Parady et al. (2021)).
Several factors affect the temporal transferability of choice models. Focusing on mode-destination choice, Fox et al. (2014) finds that improving the model specification using socioeconomic variables in the model specification improves temporal transferability, as doing so reduces the reliance on constants to explain behavior. Here, Fox et al. (2014) found that constants are the least stable model parameters. As the influence of constants diminishes with the addition of behavioral parameters, the transferability of the model improves. One of the previous studies by Parody (1977) used a before and after methodology to understand and validate how well disaggregate logit models predict the changes. This study focused on how mode choice among travelers changes by introducing a free bus service, subsequently increasing parking fees, and implementing stricter parking regulations. The study finds that disaggregate modal-choice models, particularly those using detailed socioeconomic (such as gender and occupational status of the traveler) and transportation service variables (Frequency of service, Walk time to the bus stop), demonstrated much better predictive accuracy in forecasting shifts in travel modes resulting from transportation system changes.
Using a similar before and after approach by Parody (1977), Train (1978) examined the predictive accuracy of a mode choice model focusing on the Bay Area Rapid Transit (BART) system opened in San Francisco. The study compared actual post-BART mode shares with pre-BART predictions and analyzed parameter stability. It found that transit use, especially BART with walk access, was overestimated. The study tested non-IIA models (Maximum, Log-sum) to evaluate whether IIA assumption failure caused this transit overprediction. These models also overestimated transit use; the study concluded that although the IIA assumption might slightly contribute to overprediction, other issues, like unique BART-specific attributes and inaccurate walk time data, were the primary contributors to the overprediction.

3. Data

The study area for this research is the Metropolitan Region of Amsterdam (MRA), with a focus on three trip purposes: (a) Work, (b) Education, and (c) Home-Based Maintenance (HBM) trips. Considering the data availability for the MRA region, the forecast horizon is a 5-year short-term forecasting horizon (2018-2022). Hence, the four datasets mentioned below are for 2018 and 2022.

Dutch National travel survey data (ODiN)
ODiN is a travel survey that tracks the travel behavior of the Dutch population provided by Statistics Netherlands (CBS). Participants are required to record their daily travel details, including destinations, purposes, sociodemographic characteristics, activity duration, and other trip details, for one specific day each year (DANS, 2024). The arrival points of the trips are available at the 4-digit postal code (PC4) level.

Employment data
This data represents the number of jobs available in the MRA at the PC4 level, provided by Research and Statistics, Amsterdam (O&S). It includes the total number of jobs in each PC4 zone municipality. Additionally, the data further categorizes jobs into various sectors, such as wholesale, retail, and other services.

Education Enrollment data
This dataset from the Dutch Ministry of Education, Cul-

ture, and Science (DUO) contains the locations of educational institutions across the Netherlands and the number of enrolled students. It covers primary, secondary, vocational, and higher education levels, corresponding to the Dutch education system.

Travel time matrix data

The travel time matrix for all PC4 zones in the Netherlands, covering private (walking, bicycle, car) and public transport modes, is calculated using the GeoDMS software by ObjectVision. Data sources include OpenStreetMap for road network details (e.g., road type, names, one-way streets) and General Transit Feed Specification (GTFS) for public transit details. Additionally, in the software, the PC4 centroids are determined using an address-weighted approach with the help of the Key Register of Addresses and Buildings (BAG) dataset containing the geolocation of all buildings and addresses in the Netherlands.

4. Methodology

4.1. Variable selection

In order to assess the impact of SC&AE on the temporal transferability of destination choice models, a utility specification for the model needs to be formulated. Hence we need to select variables to be included in the specification. The selection of the explanatory variables differs with the trip purpose. Additionally, the selection also depends on factors affecting temporal transferability and the available data. Because the highest common resolution across all four datasets mentioned in the Data Section is at the PC4 level, PC4 zones in the MRA region are considered as destination alternatives.

For including explanatory variables, in addition to the SC&AE parameter, three sets of explanatory variables are considered to be included in the MNL choice models: (a) Zonal size measures, (b) Travel impedance (c) Interaction of sociodemographic variables with impedance, and (d) SC&AE

Zonal size measures

Zonal size measures vary according to trip purpose. Typically, for work, total employment is considered, and hence here, total employment in the PC4 destination zone is considered. For HBM, employment across relevant sectors is considered. These include six sectors: Wholesale & Retail, Financial Institutions, Utilities, Government, Health & welfare, and Other services. Lastly, for education trips, total enrollment at the relevant education level is considered.

Travel Impedance

For travel impedance, across all trip purposes, log transformation of travel time between the PC4 zone is considered as it is equivalent to the power function of the impedance function of a gravity model Daly (1982).

Interaction of sociodemographic variables with impedance

As reviewed in section 2.3, adding sociodemographic characteristics of travelers in the specification improves the temporal transferability of destination choice models. The interaction of sociodemographic variables with impedance varies with the trip purpose. These interactions are included based on findings from previous studies related to the three trips within the scope of this study. Bhat et al. (1998) finds that older adults and women were more sensitive to travel impedance for work trips, while higher-income travelers were more willing to travel longer distances for work. Hence, the interactions of gender, age, and disposable household income level of travelers with travel time are included for work and HBM trips.

For education trips, previous research focusing on Dutch education, such as de Boer and Blijie (2006) (Primary education) and van Welie et al. (2013) (After Secondary education), finds highly significant interaction of socioeconomic status and migration groups of students with travel impedance. Hence, the interactions of migration and disposable household income levels of travelers with travel time are included for education trips.

SC&AE

Finally, to evaluate the impact of SC&AE on the temporal transferability of destination choice models and how it varies by trip purpose, these effects are included using a single Hansen-type accessibility index as shown in equation 1.

4.2. Data processing and Modelling Assumptions

After the selection of the variables, four datasets are processed to facilitate the estimation of the models. During the data processing, certain assumptions are made.

In the ODiN Data, trips for the relevant purposes in both years are filtered for arrivals in the 30 municipalities inside the MRA region. However, the number of PC4 zones increased in 2022 compared to 2018 as new PC4 zones were included in 2022 by the MRA administrative authorities. However, in the Employment data, the number of jobs in these newly added PC4 zones in 2022 is negligible. These zones exist only in the 2022 data and have negligible job numbers; hence, they are excluded. To ensure consistency, only the PC4 zones common between 2018 and 2022 are considered. Consequently, the data was further filtered to only include arrivals in these common PC4s.

Generally, people have only one primary location for work and education, where they spend a significant part of their day. Hence, to determine the primary arrival destination, it is assumed that the primary destination is where the person has the maximum activity duration. Following this assumption, for each traveler, only the trip with the maximum duration is retained in the case of multiple trips for trip purposes[1]

The travel time for each trip purpose is selected based on the highest used mode in the filtered trips in the ODiN data. Following this, car travel time is used for Work and HBM. For education trips, bike travel time is used. Intrazonal travel time data for various modes is unavailable in the travel time matrix. Therefore, the travel time for intrazonal trips in a destination zone is assumed to be half the travel time to the nearest neighboring zone.

The age of the traveler in the ODiN data is a continuous variable. It is converted to a categorical variable for different age groups, to include the interaction of age with the travel impedance (travel time). It includes five age groups as per the age groups used by CBS to report PC4-level statistics, removing the 0-6 age group, as ODiN data only consists of travelers aged 6 years and above. This approach avoids assuming a linear or monotonic relationship and simplifies forecasting by allowing predictions based on categories rather than continuous age values (Bhat et al., 1998). Similarly, the standard household disposable income level of travelers is recorded in the ODiN travel survey as decile groups. These decile groups are re-categorized into 5 groups to estimate the model parsimoniously. The groupings are performed by combining the median levels for each decile group reported by the CBS (Centraal Bureau voor de Statistiek (CBS), 2024) and the annual Dutch national median income reported by the OECD for 2018 and 2022 (Organisation for Economic Co-operation and Development (OECD), 2023).

Finally, the number of destination PC4 zones available to travelers depends on the availability of opportunities for the trip type, considering the compatibility between trip type and land use characteristics (Bernardin et al., 2018). This means excluding zones with no retail employment for HBM trips, zones with no employment as possible work locations, and zones with no enrollment at the relevant education level as possible locations for education trips. Consequently, the observed trips in ODiN data were further filtered to only include arrivals in these available PC4 zones.

For education trips, the number of trips available for each education level after the primary level was very low compared with the rest of the trips. Hence, these trips were combined to form the trip purpose 'Secondary education and above.' In these trips, travelers with attained education level at Secondary (VMBO, HAVO, VWO), vocational (MBO), and higher education (HBO, WO). After processing the ODiN data and filtering further for trips arriving only in the 364 zones in the MRA region common for both years, the final number of trips for each purpose is as follows:

Table 1: Number of trips for base and transfer year, post-processing ODiN data

Trip Purpose	2018	2022
HBM	1,874	1,851
Work	2,680	1,903
Secondary education and above	364	276
Primary education	693	395

4.3.Sampling

Including all zones in the study area in the choice set is neither a realistic representation of how travelers choose destinations nor practical in terms of computational efficiency. It is widely understood that travelers do not evaluate such a large number of alternatives when selecting a destination. They tend to automatically eliminate destinations that are too far from their origin zones. Hence, sampling of destinations is required. Sampling alternatives requires making two key decisions. First, determining the appropriate choice set size, and second, selecting an adequate sampling method.

4.3.1. Determining Choice set size

Regarding the appropriate choice set size, Guevara et al. (2016) presented a method based on a Monte Carlo experiment to determine the stability of parameter estimates by their average and standard deviation. This method is suitable for various models, including the RRM, MEV logit mixture, and logit models. Specifically, Guevara et al. (2016) varied the choice set sizes ($\tilde{J}$) and for each $\tilde{J}$, sampled K times. They then estimated the mean parameter values ($\bar{\hat{\beta}}$) and their standard deviations as follows:

$$\bar{\hat{\beta}} = \frac{1}{K}\sum_{k=1}^{K}\hat{\beta}_k(\tilde{J}) \quad \text{and} \quad \hat{\sigma}_{\hat{\beta}} = \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}\left(\hat{\beta}_k - \bar{\hat{\beta}}\right)^2}$$

[1]While retaining only the trip with the maximum activity duration is valid for work and education trips, it is unsuitable for home-based maintenance (HBM) trips, where multiple destinations are possible. Initially, the HBM data was mistakenly filtered under this assumption, overlooking Pandas Biogeme's capability to handle multiple observations per traveler (PanelObs=True). This error was identified late, and it was not possible to perform all validation steps with the correct data structure at the time of submission of the final version of this paper. However, a quick informal parameter estimation showed that the parameter estimates and t-test values for most significant parameters, especially SC&AE, and model performance in both years did not differ significantly from the results presented in the Results chapter. Therefore, the general storyline remains the same.

This approach helps identify an appropriate choice set size for the given travel survey data based on the stability of the parameter estimates. In this study, the choice set size is varied from 5 to 50 destination alternatives. For each choice set size, the mean beta values are calculated across 30 iterations of the base utility specification, inclusive of SC&AE. The choice set size determined for 2018 for each trip purpose is also used when estimating the model on 2022 Data.

4.3.2. Stratified Importance Sampling

This study uses a variant of the Stratified Importance Sampling (SIS) method adapted from Bradley et al. (1998). For each origin zone, destinations are chosen based on their distance from the origin PC4 zone and the destination's attraction size relevant to the trip purpose. In this sampling method, first, destinations are arranged in terms of proximity and then two thresholds are created, D1 (20th percentile) and D2 (60th percentile). Then destinations for each origin zone are categorized into three distance bands: closer than the 20th percentile (D1), between the 20th and 60th percentiles (D2), and beyond D2. Within these distance bands, destinations are further stratified based on the median attraction size: S1 for zones between D1 and D2, and S2 for zones beyond D2. This creates a total of five strata. The destination's attraction size is varied by the trip purpose, such as total employment for work trips or retail/service employment for HBM trips. Then, for a given sample size, destinations are proportionally sampled from these five strata: 20% each from (1) zones closer than D1, then from zones lying between D1 and D2 with (2) having attraction sizes below S1 and (3) above S1. Then, in the distance band beyond D2, zones with (4) size above S2 and (5) below S2. By reflecting proximity and attractiveness, this variant of SIS ensures one possible realistic representation of destination choice sampling.

4.4. Assessing Impact of SC&AE on temporal transferability

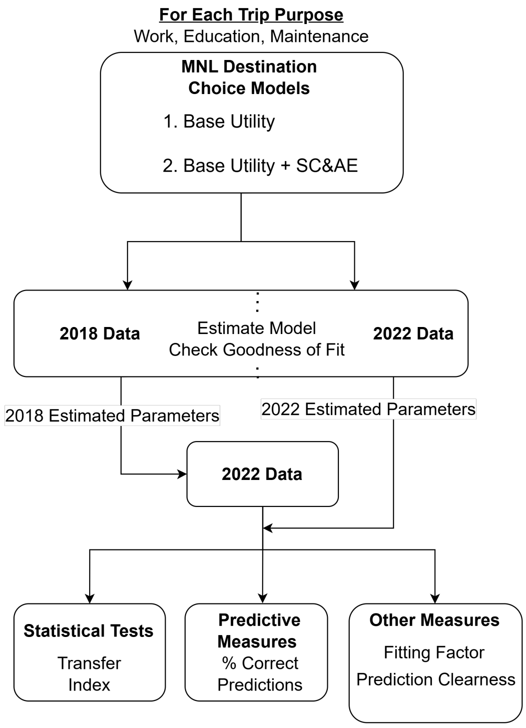


Figure 1: Methodology for Assessing Impact of SC&AE on Temporal Transferability on MNL DCMs

Figure 1 illustrates the core methodology used to assess the impact of SC&AE on the temporal transferability of destination choice models. The base utility specification includes all the variables selected in Section 4.1, except for the SC&AE parameter. Two models (one with base utility and the other inclusive of SC&AE) are estimated using data from 2018 (base year) and 2022 (forecast year) for each trip purpose. The 2018 model parameters are then applied to 2022 data and compared with models re-estimated on 2022 data. The comparison involves four measures across three categories: (1) Transfer Index, (2) Percentage of correct predictions and other measures, (3) Fitting Factor, and (4) Discriminative ability of the model.

As noted in Section 4.1, sampling in destination choice modelling is essential for a realistic behavioral representation and computational feasibility. Consequently, sampling errors are inevitable. To thoroughly understand the impact of SC&AE), this study also applies the framework illustrated in Figure 1 not only to models estimated using SIS, but also to models estimated with the same choice set size using random sampling. Random sampling is one of the most commonly used methods in destination choice modeling, in addition to various variants of SIS (Kim and Lee, 2017). This approach provides insight into how the impact of SC&AE on temporal transferability varies with the sampling method. Moreover, because random sampling is known to be an unrealistic method (assigning equal probability to all possible destination alternatives in the study area), it allows us to explore how selecting an incorrect sampling method affects the temporal transferability of DCMs, a factor not previously explored in research.
In addition to sampled models, a full-choice estimation was performed for the three trip purposes, considering all available alternatives without any sampling. This was done to establish an unbiased benchmark, free from sampling-induced errors, against which the performance of sampled models could be compared. By eliminating sampling biases, the full-choice estimation offers a clearer view of the true effects of SC&AE and provides a baseline for assessing how different sampling

methods influence parameter estimates and temporal transferability.

Selected Performance Indicators:
The parameter estimation of discrete choice models is based on the Maximum likelihood principle, which is a statistical method to estimate model parameters by finding values that maximize the likelihood of observing the given data (Bunch, 1987). Hence, log-likelihoods are commonly used to assess the explanatory power of a specified discrete choice model. Therefore, the first test of transferability is a statistical test (Transfer Index), whose calculation is also based on the log-likelihoods of models estimated across the forecast horizon. However, because these models are used as predictive tools, it is important to quantify or translate how performances on statistical tests translate to the practical use of the model. Therefore, the second category chosen to assess the impact is Predictive measure such as % Correct predictions. However, discrete choice models are probabilistic models, not deterministic (Hauser, 1978). Discrete choice models, such as the Multinomial Logit, predict the probability of choosing an alternative, not the actual choice itself. Hence, other measures, such as the Fitting factor and Prediction clarity, are required to assess the quality of probability predictions.

Transfer Index (TI)
Developed by Koppelman and Wilmot (1982), the Transfer Index (TI) is a statistical test that tests the extent of model transferability (Fox et al., 2014). It quantifies the temporal transferability of a model by comparing the predictive accuracy of a transferred model (using base year parameters) with that of a locally estimated model (re-estimated using transfer year data). It is calculated as the ratio of the log-likelihood (LL) improvement gained by re-estimation to the loss in LL when using transferred parameters, both relative to a simplistic reference model. A TI of 1 indicates perfect transferability, whereas values below 1 reflect reduced performance, with negative values indicating that the transferred model performs worse than the reference model. In this study's context, the 2018 ODiN Data is the base year sample, and the 2022 ODiN Data is the transfer year sample. The disaggregate equivalent of a gravity model is used as a simplistic reference model, with its destination utility specification having only travel impedance and size measure.

% Correct Predictions:
As a predictive test, it evaluates the accuracy of the model by calculating the ratio of correct predictions to the total number of observations expressed as a percentage. The alternative to which the model assigns the highest probability among all alternatives in the choice set is the predicted choice (Parady et al., 2021).

Prediction Clearness:
One of the key limitations of the above % Correct Predictions is its inability to account for the model's discriminative ability in its evaluation. Precision Clearness addresses the limitations of percentage correct predictions by evaluating the model's ability to distinguish the observed choice clearly. It includes three measures based on a probability threshold: % Clearly Right (CR), where the model assigns the observed choice a probability above the threshold; % Clearly Wrong (CW), where the model assigns any alternative other than the observed choice above the threshold; and % Unclear (UC), where no choice is assigned a probability above the threshold (De Luca and Cantarella, 2009).

Fitting Factor (FF):
Considering the sample size, the FF measures the average probability that a model assigns to the observed choice. It has an upper bound of one, indicating that on average, the model assigns a probability of one to the observed choice; hence, it perfectly forecasts all choices in the sample (De Luca and Cantarella, 2009).

5. Results

Table 2: Number of available destination alternative and choice set size across trip purpose

Trip Purpose	Number of available PC4 zone destinations		Choice set size
	2018	2022	
HBM	350	355	40
Work	363	363	45
Secondary and above Education	125	129	45
Primary Education	278	282	278

Table 2 presents the number of available PC4 destination zones with nonzero zonal size measures and the corresponding choice set chosen for each trip purpose. Initially, for trips, the model was estimated on the full choice set (including all available destinations), before estimating the parameters for the corresponding choice set size determined for sampling. As the full choice set would include all destinations, this was done to check whether the parameter estimate was statistically significant for each trip purpose. Moreover, the SC&AE parameter remained relatively the same, with similar statistical significance across the full and sampled choice sets. Hence, this approach helped save time by avoiding determining the choice set size for sampling when the SC&AE parameter was found to be statistically insignificant for the full choice set.
For primary education, the SC&AE parameter was found

to be statistically insignificant in the 2018 data for the full choice set. Hence, the method of determining the choice set size was not used for primary education trips.

The open-source Python package Pandas Biogeme is used for parameter estimation.

5.1. SC&AE Parameter estimates across trip purposes

Table 3: SC&AE Parameter estimates for 2018 and 2022 across trip purposes

Trip Purpose	SC&AE Parameter estimates (Robust t-test values)	
	2018	2022
HBM	-1.60 (-8.56)	-1.77 (-10.55)
Work	-0.67 (-7.09)	-0.76 (-7.15)
Secondary and above Education	0.72 (4.61)	0.48 (2.51)
Primary Education	-0.22 (-1.23)	NA

Table 3 presents the SC&AE parameter estimates for various trip purposes, for the corresponding choice set sizes shown in Table 2. Except for primary education, the SC&AE parameter for all trip purposes is statistically significant at the standard 5% level. Given the negative SC&AE parameters for HBM and Work, the results suggest that spatial competition dominates the agglomeration effect for both these trip purposes in 2018 and 2022. These results are consistent with the findings of previous studies for HBM (Bhat et al. (1998); Bernardin et al. (2009)) and for work Ho and Hensher (2016).
On the other hand, for secondary and above education trips, the estimated SC&AE parameter has a positive sign, suggesting that agglomeration effects dominate spatial competition. These results contradict the findings of a previous study explaining university location choices among high school graduates in the Netherlands by Sá et al. (2004), where they found a dominating competition effect (negative parameter). However, their model focused solely on choices for destination choices for one level immediately after secondary education, where spatial competition may play a more significant role since only one education level is included. This study included trips for both secondary and multiple higher education levels, thus including destinations with opportunities for students to have a broader range of options to continue education nearby. This proximity to further educational opportunities can make zones with more secondary and post-secondary institutions attractive, hence a positive value for the dominating agglomeration effect. Finally, for primary education trips, the estimated SC&AE parameter in 2018 is statistically insignificant at the 5% confidence interval. This suggests that SC&AE do not play a significant role in explaining destination choices for primary education trips. Given the statistical insignificance in 2018, it does not allow for analyzing the impact of SC&AE on the temporal transferability for this trip purpose. Thus, the analysis was cut short at this point without estimating the parameters for 2022.
Comparing the value of SC&AE parameter estimates for HBM and Work, HBM trips had a more negative parameter (-1.60 to -1.77) than Work (-0.67 to -0.76), indicating that SC&AE has a stronger influence on destination choices for HBM trips than for work trips. This seems reasonable from the perspective of the ease of switching to alternative destinations. HBM trips, which include shopping and personal errands, often involve destinations that are closer substitutes (e.g., multiple grocery stores or service centers within a short distance), intensifying competition. On the other hand, work trips generally involve more specialized destinations (e.g., offices or job locations), where alternatives are more limited, resulting in weaker spatial competition. Additionally, travelers performing HBM trips have a very high level of ease of switching, as they have no mandate or commitment to stick to a specific shopping or service destination to perform maintenance activities. On the other hand, employment is a long-term decision with a longer commitment period, and switching jobs is not as easy as switching destinations to perform maintenance activities. Hence, it makes sense that the spatial distribution of opportunities has a lower influence on destination choices for work trips than for HBM trips.

5.2. Impact of SC&AE on temporal transferability across trip purposes

5.2.1. Home-based Maintenance trips

Table 4: HBM trips: Performance Comparison on various indicators Full vs. Sampled choice sets (RS and SIS)

HBM 2022 (using 2018 estimated parameters)				
Full Choice Set				
Indicators	Base Utility	Base Utility + SCAE	% Improved	Absolute gain in TI value
TI	-0.44	0.82	286.36%	1.26
Fitting Factor	0.2016	0.2068	2.59%	-
% Correct Prediction	36.09%	36.14%	0.14%	-
40 Alt Randomly Sampled				
TI	-1.21	0.32	126.45%	1.53
Fitting Factor	0.575	0.580	1%	-
% Correct Prediction	70.50%	71.20%	0.99%	-
40 Alt SIS				
TI	0.04	0.63	*	0.59
Fitting Factor	0.538	0.545	1.39%	-
% Correct Prediction	66.7%	66.6%	-0.15%	-

*This figure was unrealistically high due to the low denominator (0.04). Hence it was not reported.

Table 4 compares the performance of full-choice sets with models using 40 randomly sampled and stratified importance sampled (SIS) alternatives for HBM trips in 2022. The models, estimated on 2022 data using 2018 parameters, are evaluated across multiple performance indicators: Transfer Index (TI), Fitting Factor, and Percentage of Correct Predictions. This comparison highlights the impact of the SC&AE parameter and the differences between the sampling methods.

For the full-choice set, including SC&AE significantly improves the TI value from -0.44 (Base Utility) to 0.82. This suggests that without the SC&AE parameter, the model performs worse than the simple disaggregate equivalent of a gravity model estimated on 2022 data indicated by the negative sign. But when the SC&AE parameter is included, the model retains 82% of the performance gain achieved by re-estimating parameters on 2022 data. This demonstrates a significant positive impact of SC&AE. However, other metrics, such as Fitting Factor (2.59% improvement) and Percentage of Correct Predictions (0.14% improvement), show much smaller gains, suggesting that the TI may exaggerate the impact of SC&AE.

Using random sampling, for the base utility specification, the TI value (-1.21) is highly negative and nearly three times lower than the base utility specification of the full choice set (-0.44). This highlights that random sampling leads to reduced temporal transferability.

Interestingly, with SIS, the base Utility model achieves a positive TI value (0.04), unlike full-choice or random sampling, where the TI for base utility is negative. This indicates that SIS minimizes sampling biases, enabling the model estimated on 2018 parameters to outperform the 2022 disaggregate equivalent gravity model (reference model), even without SC&AE. This increased initial performance by the base utility specification also means that the impact of SC&AE on temporal transferability in terms of absolute gain in TI value (0.59) is reduced compared with Random sampling (1.53) and the full choice set (1.26). However, in terms of the achieved TI value after including SC&AE, compared with random sampling (0.32), SIS allows the model to reach a much higher TI value (0.63). The Percentage of Correct Predictions slightly decreases (-0.15%), However, this metric does not account for model's discriminative ability in its evaluation. Observing the % Clearly Right across thresholds 40% to 90% in figure 2 confirms the consistent positive but limited impact of SC&AE.

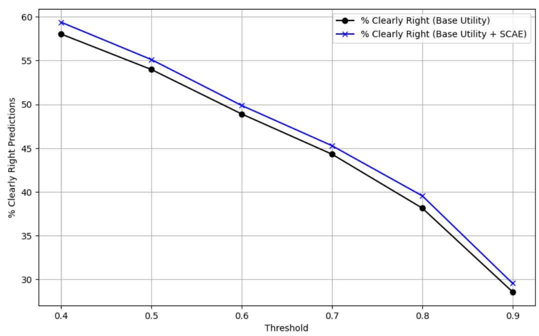


Figure 2: % Clearly Right for HBM Trips in 2022 Using 2018 Estimated Parameters (Stratified Importance Sampling)

5.2.2. Work

Table 5: Work trips: Performance Comparison on various indicators Full vs. Sampled choice sets (RS and SIS)

Work 2022 (using 2018 estimated parameters)				
Full Choice Set				
Indicators	Base Utility	Base Utility + SCAE	% Improved	Absolute gain in TI value
TI	0.54	0.73	35.2%	0.19
Fitting Factor	0.0263	0.0272	3.42%	-
% Correct Prediction	7.93%	7.99%	0.75%	-
45 Alt Randomly Sampled				
TI	0.437	0.662	52.01%	0.22
Fitting Factor	0.147	0.152	3.61%	-
% Correct Prediction	26.06%	26.59%	2.03%	-
45 Alt SIS				
TI	0.48	0.71	47.91%	0.23
Fitting Factor	0.138	0.143	3.62%	-
% Correct Prediction	25.28%	25.59%	1.23%	-

Table 5 presents the results for work trips in 2022 using the 2018 estimated parameters, across full-choice, random sampling (RS), and stratified importance sampling (SIS) methods. Compared to HBM trips, including SC&AE in work trips exhibits smaller improvements in the Transfer Index (TI) and other indicators, reflecting weaker SC&AE impacts on temporal transferability for this trip purpose.
For the full-choice set, including SC&AE improves the TI from 0.54 to 0.73, representing a 35.2% increase. While this demonstrates a positive impact of SC&AE, the improvement is less dramatic than the jump observed for HBM trips (-0.44 to 0.82). Other indicators, such as the Fitting Factor and Percentage of Correct Predictions, also showed limited improvements (3.42% and 0.75%, respectively), further supporting the limited impact of SC&AE on the temporal transferability for work trips.
In the random sampling approach, TI increased from 0.437 to 0.662 with SC&AE, showing better transferability than that observed for HBM trips, where TI remained highly negative (-1.21 to 0.32).
With SIS sampling, the TI improves from 0.48 to 0.71, a 0.23 increase in absolute value, less than that for HBM using SIS (0.59). Notably, the Percentage of Correct Predictions shows a gain of 1.23% with SC&AE, in contrast to the slight decline (-0.15%) for HBM trips.
Interestingly, although a similar trend of exaggerated impact by TI and limited impact on other indicators is also observed here, unlike HBM trips where SC&AE had the most significant impact on TI, the improvements for work trips were more balanced across TI, Fitting Factor, and Correct Predictions. This suggests that while SC&AE enhances transferability for work trips, its influence is less dominant than that of HBM trips. This is also reflected in the improvements due to SC&AE in % Clearly Right predictions in figure 3. For HBM trips (Figure 2), the gap between the Base Utility and Base Utility + SC&AE specifications is more noticeable for HBM trips, indicating a stronger impact of SC&AE for HBM trips. In contrast, for work trips, the two curves get closer with the increase in the thresholds.

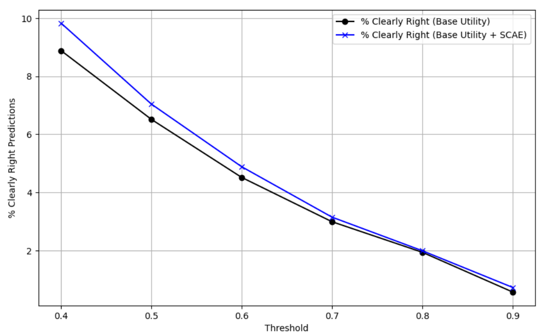


Figure 3: % Clearly Right for Work Trips in 2022 Using 2018 Estimated Parameters (Stratified Importance Sampling)

5.2.3. Secondary and Above Education

Table 6: Secondary and Above Education trips: Performance Comparison on various indicators Full vs Sampled choice sets (RS and SIS)

Secondary and Above Education 2022 (using 2018 estimated parameters)				
Full Choice Set				
Indicators	Base Utility	Base Utility + SCAE	% Improved	Absolute gain in TI value
TI	0.30	0.43	43.33%	0.13
Fitting Factor	0.0657	0.0632	-3.81%	-
% Correct Prediction	14.86%	12.68%	-14.7%	-
45 Alt Randomly Sampled				
TI	-0.56	-0.05	91.07%	0.51
Fitting Factor	0.143	0.139	-2.88%	-
% Correct Prediction	33.70%	28.26%	-16.1%	-
45 Alt SIS				
TI	-0.02	0.28	*	0.30
Fitting Factor	0.144	0.141	-1.81%	-
% Correct Prediction	32.25%	28.62%	-11.3%	-

*This figure was unrealistically high due to the low denominator (0.02). Hence, it was not reported.

Table 6 presents the results for secondary and above education trips in 2022 using the 2018 estimated parameters across full-choice, random sampling (RS), and stratified importance sampling (SIS) methods. Compared to HBM and work trips, the improvements due to SC&AE for education trips are smaller on the TI metric and negative on other metrics. For the full-choice set, the inclusion of SC&AE improves the TI from 0.3 to 0.43, a 43.33% increase, similar to work trips. However, the absolute gains are even lower. This reflects a weaker impact of SC&AE for secondary and above education trips. For random sampling, the TI improvement increases from -0.56 to -0.05, a good absolute increase of 0.51, yet it is still in the negative. and shows notable declines in the Fitting Factor (-2.88%) and Percentage of Correct Predictions (-16.14%), indicating that with random sampling, including SC&AE leads to the model performing poorly for secondary and above education trips. These results suggest that SC&AE has negligible relevance for secondary and above education trips.

With SIS sampling, including SC&AE, the TI improves significantly from -0.02 to 0.28. This suggests that the impact of SC&AE on temporal transferability improves with a better sampling method, a trend observed in HBM and Work trips too. This is also observed when comparing the performance on % Clearly Right plot of Random Sampling (figure 4) and SIS (figure 5).

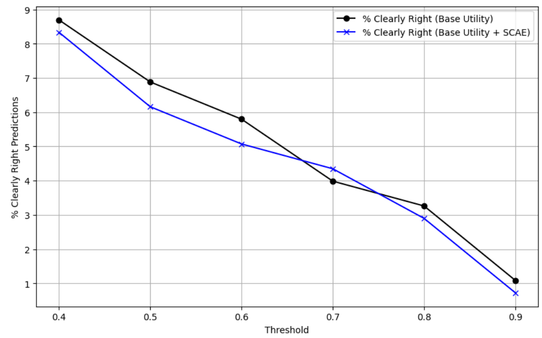


Figure 4: % Clearly Right for Secondary and Above Education Trips in 2022 Using 2018 Estimated Parameters (Random Sampling)

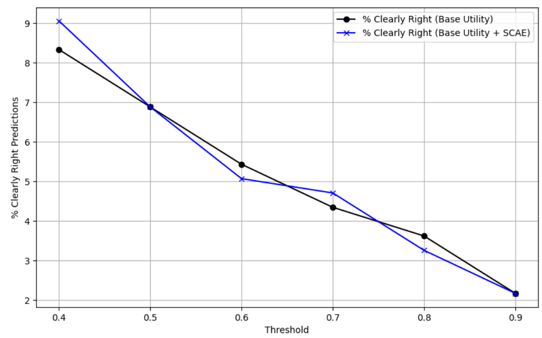


Figure 5: % Clearly Right for Secondary and Above Education Trips in 2022 Using 2018 Estimated Parameters (Stratified Importance Sampling)

Overall, unlike HBM and work trips, secondary and above education trips show declines in key metrics like Fitting Factor, % Correct Predictions and Clarity Analysis. with the inclusion of SC&AE, especially for sampled sets. This suggests that the SC&AE parameter may be less relevant for secondary and above education trips.

5.3.Source of Transfer Index's Exaggerated Impact

The results presented in section 5.2 suggest that the Transfer Index generally tends to exaggerate the impact of SC&AE on the temporal transferability of models across all three trip purposes. This section provides insights into the source of this exaggeration through a comparative analysis of the LLs of the models used in the TI calculation for randomly sampled HBM trips, as presented in Table 7. Additionally, in this section, we also provide insight into why, for secondary and above education trips, the inclusion of SC&AE shows a positive impact on TI, but fails to translate into consistent improvements across other performance measures.

Table 7: TI Comparative analysis: Source of exaggerated Impact (HBM, Randomly Sampled)

	GM	Base Utility	Base Utility + SCAE
	2022		
Final LL	-2016.42	-2003.18	-1970.48
Gain over GM(2022)	—	13.24	45.94
2022 (using 2018 parameters)			
Final LL		-2032.40	-2001.62
Gain over GM(2022)	—	-15.98	14.81
% Change in gain		-220.65%	-67.77%
TI		**-1.21**	**0.32**

Null LL: -6828.12

As shown in Table 7, adding the SC&AE parameter for HBM trips results in a significant LL gain of 45.94 over the GM model when re-estimated for 2022. Using the 2018 parameters, the SC&AE model still achieved a gain of 14.81 LL, representing only a 67.77% reduction from the maximum achievable gain of 45.94. This explains the relatively high TI value of 0.32, indicating that the model using 2018 parameters retains 32% of the performance achieved by the re-estimated model.

Conversely, the base utility specified model, with 2018 parameters, performs slightly worse than the GM model, but the difference is small, only -15.98 LL. Although this difference is minor, it appears exaggerated compared to the minimal 13.24 LL gain achieved by re-estimating the Base Utility model, leading to a percentage change in the gain of -220.65%. This reflects how TI is similarly calculated by comparing the loss in gains to the potential gains from re-estimation (-15.98/13.24). Hence, focusing solely on TI performance can exaggerate the impact of the SC&AE parameter.

Table 8: TI Comparative analysis: Source of Inconsistent Impact (Secondary and Above Education Location Choice, SIS)

	GM	Base Utility	Base Utility +SCAE
	2022		
Final LL	-770.79	-763.82	-760.24
Gain over GM(2022)	—	6.98	10.56
2022 (Using 2018 parameters)			
Final LL		-770.91	-767.83
Gain over GM(2022)	—	-0.11	2.97
% Change in gain		-101.62%	-71.89%
TI		**-0.02**	**0.28**

Null LL: -1050.64

To explain why the inclusion of SC&AE has a positive impact on TI for secondary and above education trips but fails to improve other performance measures, we need to compare the LLs of models used in the TI metric for HBM trips (Random Sampling) presented in Table 7 with that for secondary and above education trips (table 8) . For HBM trips, including SC&AE, achieves a TI value of 0.32, with absolute LL gains of 45.94 for re-estimation in 2022 and 14.81 using 2018 parameters, over the reference GM model. In contrast, as seen in Table 8, for secondary education trips, the achievable TI value with SC&AE is similar at 0.28, but the absolute LL gains are much smaller: 10.56 from re-estimation and 2.97 using 2018 parameters over the GM model.

Clearly, the SC&AE parameter has a much stronger impact on improving model performance for HBM trips, as evidenced by the larger LL gains. However, because TI compares the ratio of these gains rather than their absolute values, it presents the maximum achievable transferability for SC&AE in both cases at similar levels. This explains why SC&AE had a positive impact on all indicators for HBM trips, yet showed a negative impact on these same indicators for secondary and above education trips, despite demonstrating a positive effect on the TI metric. The key issue lies in the relatively low absolute gains in the LL for secondary education trips.

6. Discussions

This study is the first to validate the theory of Spatial Competition and Agglomeration Effects (SC&AE) beyond the single-period goodness-of-fit measures. Consistent with previous research, it finds statistically significant negative SC&AE parameters for home-based maintenance (HBM) and work trips in 2018 and 2022, indicating a dominant spatial competition. HBM trips exhibited more negative SC&AE parameters (-1.60 to -1.77) than work trips (-0.67 to -0.76), reflecting a stronger influence on destination choices due to easily substitutable destinations and high ease of switching. In contrast, work trips involve destinations with fewer alternatives nearby and longer-term commitments, resulting in a lesser influence of spatial competition.

For secondary and higher education trips, the positive SC&AE parameter suggests that agglomeration effects dominate spatial competition. Unlike Sá et al. (2004), who found dominating spatial competition in university choices among high school graduates in the Netherlands by focusing solely on one post-secondary education level, this study included both secondary and multiple higher education levels. Hence, the broader range of nearby educational options to continue education makes areas with a higher number of institutions more attractive, resulting in a positive value that reflects dominant agglomeration effects. For primary education trips, the SC&AE parameter in 2018 was statistically insignificant, indicating a negligible influence on destination choices. This could be due to young children having the highest commitment period and the lowest flexibility to switch schools. Hence, the spatial distribution of primary school opportunities has little effect on their location choices.

Additionally, this study extends beyond evaluating the explanatory power of SC&AE in a single context by exploring its contribution to the temporal transferability of destination choice models. The findings indicate that SC&AE has a positive but limited impact on temporal transferability, varying by trip purpose: highest for home-based maintenance (HBM) trips, followed by work trips, and inconsistent for secondary & higher education (positive on the Transfer Index but negative on rest of the indicators) .

Notably, the Transfer Index (TI) performance metric tends to exaggerate SC&AE's positive impact when viewed in isolation; a closer examination of log-likelihoods and other performance indicators (such as fitting factor, percentage of correct predictions, and clearness of predictions) confirms that the actual impact is limited. This exaggeration occurs because TI relies solely on the ratio of gains in log-likelihoods (LL), which can misrepresent models with small absolute gains. For example, a model with an LL gain ratio $\frac{1}{2}$ and another with a gain ratio of $\frac{50}{100}$ will have the same TI value of 0.5. But clearly, the second model is much better and will perform positively on other indicators, while the model with LL gain ratio $\frac{1}{2}$ will perform poorly on other indicators. Such a case was also discussed in Section 5.3.

Therefore, TI values should always be presented alongside other performance measures or at least be accompanied by a comparative analysis of the log-likelihood values used in their calculation.

Observing the trend of varying impacts of SC&AE on temporal transferability by trip purpose, the impact of SC&AE on temporal transferability decreases with decreasing traveler autonomy and ease of switching destinations: it is highest for HBM trips, where travelers have high autonomy and flexibility, less so for work trips due to longer commitment periods, and inconsistent (positive on the Transfer Index but negative on rest of the indicators) for secondary and above level education trips, where travelers are effectively committed to institutions until they complete their education. Considering the low amount of trips for secondary and above education trips compared to other trips, these inconsistent results for secondary and above education trips should be considered inconclusive. However, considering autonomy and ease of switching, the impact is likely to be limited, more so than for the HBM and work trips.

For primary education trips, autonomy and flexibility are the lowest, thus explaining the statistically insignificant estimated SC&AE parameter for primary education trips.

Moreover, the choice of sampling method affects the temporal transferability of destination choice models and, thus, the impact of SC&AE on temporal transferability too. Models using Stratified Importance Sampling (SIS) show higher TI values than those using Random Sampling (RS), with and without the SC&AE parameter. While the initial performance boost from SIS reduces the absolute gain in the TI value from including SC&AE compared to RS, SIS allows models to achieve higher overall TI values after including SC&AE.

7. Limitations

This study's five-year forecasting horizon (2018–2022) was limited by data availability from the ODiN travel survey. A longer period (7–10 years) would better capture SC&AE's long-term effects as destination choices evolve gradually, and policymakers prefer extended forecast horizons. Additionally, the overlap with the COVID-19 pandemic likely altered travel behaviors, so the findings may reflect pandemic-influenced patterns rather than ones in normal conditions. Without comparing normal and pandemic-like rare conditions, our understanding of SC&AE's impact on temporal transferability under different scenarios is limited.

The travel time used for each trip is based on the highest used mode by travelers for each trip, which, while practical, overlooks the availability of other modes. A weighted average approach was tested, resulting in unrealistically high travel times, so the highest-frequency mode was used. This limitation affects interpretation because SC&AE and other model parameters interact with travel time across all trip purposes.

Due to the low amount of observed data for each education level after primary education, this study combined travelers across different education stages into a single destination choice model for secondary and higher education trips. Ideally, separate models for each education level should have been developed to capture the unique factors influencing destination choices specific to education levels. This limitation fails to capture the important differences in travel behavior among students at various stages, undermining the results. For example, the factors influencing destination choices for secondary school students can differ significantly from those affecting university students. Moreover, the low amount of observed data renders the impact of SC&AE on temporal transferability as inconclusive for secondary and above education trips. These two factors limit our understanding of how SC&AE impacts the temporal transferability of destination choice models for education trips. To overcome this, richer data regarding traveler trip observations and information on the institutions, such as education level and course offered, should be used. As these factors play a major role in explaining destination choices for education trips, as seen in previous research. Consequently, using this richer data, separate models for each education level should be developed to capture the unique factors influencing destination choices specific to education levels.

Finally, due to time constraints, this study did not apply the ODiN survey's weighting factors, which adjust the data to represent the broader Dutch population by compensating for response bias and sample selectivity. Not using these weights limits the study's representativeness and may skew the results, making population-level inferences less accurate by reflecting sample-level trends rather than on a population level.

8. Conclusions

Overall, SC&AE has a limited but positive impact on the temporal transferability of the MNL destination-choice models. This impact decreases with decreasing traveler autonomy and ease of switching to alternate destinations. Given the minimal effort required to include these effects in an MNL model because it reuses existing information such as zonal size measures and travel impedance, SC&AE provides technically "free" robust log-likelihood gains.

From a policy perspective, including SC&AE into destination choice models enhances their effectiveness as predictive tools by improving their temporal transferability. Models that include SC&AE maintain predictive accuracy over time, which is particularly valuable for scenarios where travelers have significant autonomy and flexibility, such as discretionary activities such as shopping and maintenance trips. This added robustness stems from SC&AE's ability to address two fundamental flaws of Multinomial Logit (MNL) models that limit their behavioral accuracy in representing travelers' destination choices: (1) the Independence of Irrelevant Alternatives (IIA) assumption, and (2) neglecting the influence of the spatial distribution of opportunities. By including information about all alternative destinations, SC&AE tackles both issues, allowing MNL models to become more behaviorally representative whilst retaining their computational simplicity.

The results of this study have broader implications for transport modeling. To overcome the limitations of MNL models, researchers have often relied on more complex disaggregate models, which are computationally intensive and often impractical for large datasets. However, this study demonstrates that simpler models such as MNL can overcome their flaws by integrating theories such as SC&AE, thereby improving behavioral representation while retaining computational efficiency. By focusing solely on whether a model explains or predicts behavior well, we may have been asking the wrong questions. Instead, we should ask, "Is my model an accurate representation of the system it is supposed to represent?" By addressing this fundamental question and bringing models closer to accurately reflecting the system, we automatically enhance their explanatory and predictive capabilities. This research shows that we do not necessarily need to rely on more complex models; there is another simpler way: using theories to enhance simpler models. When these simple models reach their limit on how much they can be improved using behavioral theories, which they eventually will, this approach might allow another pathway to improve more complex models, making them more computationally feasible and data-efficient without relying primarily on advancements in computational science.

To achieve this, we need to look beyond the transport domain and draw insights from related fields, such as psychology or other behavioral sciences. It may be time for transport modelers to look beyond their transport domain and integrate psychological theories to make transport models a better representation of how travelers make choices. The results of this study are certainly encouraging.

9. Recommendations and Future Research

The findings suggest that including SC&AE in destination choice models requires minimal effort and leverages existing data, such as zonal size measures and travel impedance, providing "free" robust log-likelihood gains. Therefore, the main recommendation is to first account for the limitations mentioned above and perform the required calibrations. and then include SC&AE parameters in destination choice models, especially for trips where travellers have significant autonomy and flexibility in choosing destinations, such as shopping and maintenance activities.

The validation methodology used in this study can be adapted into a comprehensive experimental scenario analysis framework to test other theories or model enhancements. Should future high-impact event like pandemic occur, the insights from this study will enable researchers to better assess model performance under such conditions by experimenting with different specifications and theories. As more ODiN data becomes available, researchers can adjust the forecasting horizon to examine model transferability over different periods, allowing for thorough comparative analyses and building confidence among transport planners and stakeholders.

For future research, it is necessary to address mode assumptions by using travel times inclusive of all available transportation modes. Exploring the use of the mode log sum is promising, as it captures the combined accessibility effects and reflects travelers' mode choices more accurately. Adapting the SC&AE formulation to integrate the mode log sum could enhance the model's representation of accessibility.

Additionally, this framework can be utilized to assess the impacts of various enhancements on subsequent choice models within a comprehensive transport modeling system, which often includes multiple interconnected models, such as route choice, mode choice, and route assignment. By examining how improvements in the destination choice model influence other components, we can better understand the cascading effects throughout the modeling framework and how they contribute to overall predictive performance.

Further research could also investigate more complex formulations of SC&AE, such as defining them using two accessibility factors as introduced by Bernardin et al. (2009). Although this approach may require more effort in data collection and processing, potentially diminishing the appealing trade-off of minimal effort for log-likelihood gains, it could offer enhanced benefits that extend to subsequent choice models. Combining this advanced formulation with previous suggestions may provide a deeper understanding of how such enhancements affect the entire transport modeling framework, ultimately leading to more accurate and robust models for policy analysis and decision-making.

In the longer term, to improve predictive performance, research should focus on ensuring that models accurately represent the decision-making systems they are trying to represent, rather than solely focusing on improving explanatory or predictive abilities. This can be achieved by developing and including theories to address fundamental flaws. Drawing from relevant fields, such as psychology or other behavioral science, to further refine models, making them more representative of actual decision-making processes. Such an approach would help use simple models and data more efficiently to build robust models. This approach may even provide insight into how to make complex models more computationally feasible. By prioritizing behavioral representativeness, models naturally achieve better explanatory and predictive performance.

Acknowledgements

This study was conducted in collaboration with TNO, the Dutch Organization for Applied Scientific Research, in an effort to develop a destination choice model for their Activity Based Modeling framework. The author thanks TNO and the supervisor from their SUMS department, Dr. Ir. Han Zhou, for their support during the course of this research. Additionally, comments from Dr. Ir. Gonçalo Correia and Dr. Ir. Maarten Kroesen from TU Delft have played a crucial role in improving the content of this paper.

References

Bernardin, V., Chen, J., and daniels, c. (2018). How-to: Model destination choice. Technical report, United States Department of Transportation Federal Highway Administration.

Bernardin, V. L., Koppelman, F., and Boyce, D. (2009). Enhanced destination choice models incorporating agglomeration related to trip chaining while controlling for spatial competition. *Transportation Research Record: Journal of the Transportation Research Board*, 2132:143–151.

Bhat, C., Govindarajan, A., and Pulugurta, V. (1998). Disaggregate attraction-end choice modeling formulation and empirical analysis. *Transportation Research Record*.

Bradley, M. A., Metro, P., Bowman, J. L., and Systematics, C. (1998). A system of activity-based models for portland, oregon. Technical report, U.S. Department of Transportation and Environmental Protection Agency. USDOT Report DOT-T-99-02. Produced for the Travel Model Improvement Program of the U.S. Department of Transportation and Environmental Protection Agency, Washington, D.C.

Bunch, D. S. (1987). Maximum likelihood estimation of probabilistic choice models. *SIAM Journal on Scientific and Statistical Computing*, 8.

Centraal Bureau voor de Statistiek (CBS) (2024). Statline - dataset 83932ned. Accessed: October 29, 2024.

Daly, A. (1982). Estimating choice models containing attraction variables. *Transportation Research Part B: Methodological*, 16:5–15.

DANS (2024). Onderzoek onderweg in nederland. Accessed: 2024-06-02.

de Boer, E. and Blijie, B. (2006). Modelling school choice in primary education: An aid in school location planning - an exercise with the city of zwijndrecht schools. Report from Delft University of Technology, Department of Transport and Planning.

de Dios Ortúzar, J. and Willumsen, L. G. (2011). *Modelling Transport*. Wiley.

De Luca, S. and Cantarella, G. E. (2009). Validation and comparison of choice models. In Saleh, W. and Sammer, G., editors, *Travel Demand Management and Road User Pricing: Success, Failure and Feasibility*, pages 37–58. Ashgate Publications.

Fotheringham, A. S. (1985). Spatial competition and agglomeration in urban modelling. *Environment and Planning A: Economy and Space*, 17.

Fox, J., Daly, A., Hess, S., and Miller, E. (2014). Temporal transferability of models of mode-destination choice for the greater toronto and hamilton area. *Journal of Transport and Land Use*, 7:41–62.

Guevara, C. A., Chorus, C. G., and Ben-Akiva, M. E. (2016). Sampling of alternatives in random regret minimization models. *Transportation Science*, 50:306–321.

Hauser, J. R. (1978). Testing the accuracy, usefulness, and significance of probabilistic choice models: An information-theoretic approach. *Operations Research*, 26.

Ho, C. Q. and Hensher, D. A. (2016). A workplace choice model accounting for spatial competition and agglomeration effects. *Journal of Transport Geography*, 51:193–203.

Kim, J. and Lee, S. (2017). Comparative analysis of traveler destination choice models by method of sampling alternatives. *Transportation Planning and Technology*, 40:465–478.

Koppelman, F. S. and Wilmot, C. G. (1982). Transferability analysis of disaggregate choice models. *Transportation Research Record*, 895.

Organisation for Economic Co-operation and Development (OECD) (2023). Data explorer - oecd wise database: Water indicators. Accessed: October 29, 2024.

Parady, G., Ory, D., and Walker, J. (2021). The overreliance on statistical goodness-of-fit and under-reliance on model validation in discrete choice models: A review of validation practices in the transportation academic literature. *Journal of Choice Modelling*, 38:100257.

Parody, T. E. (1977). Analysis of predictive qualities of disaggregate modal-choice models. *Transportation Research Record*.

Sá, C., Florax, R. J. G. M., and Rietveld, P. (2004). Determinants of the regional demand for higher education in the netherlands: A gravity model approach. *Regional Studies*, 38:375–392.

Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46:234–240.

Train, K. (1978). A validation test of a disaggregate mode choice model. *Transportation Research*, 12.

van Welie, L., Hartog, J., and Cornelisz, I. (2013). Free school choice and the educational achievement gap. *Journal of School Choice*, 7:260–291.

Appendix

Table 9: Estimated Betas, HBM (2018, 2022 ODiN), Full vs 40 Sampled Choicesets (RS and SIS)

Estimated Betas, HBM (2018, 2022 ODiN), Full vs 40 Alt Choice set (RS and SIS)							
		2018			**2022**		
		Full	**40 RS**	**40 SIS**	**Full**	**40 RS**	**40 SIS**
Note	Name	Parameter Values (t-test values)			Parameter Values (t-test values)		
Parameters for logarithmic values	Jobs	0.98 (26.42)	1.08 (22.92)	1.11 (23.24)	0.94 (24.91)	1.02 (19.73)	1.01 (21.75)
	SCAE	-1.40 (-10.46)	-1.35 (-7.17)	-1.60 (-8.56)	-1.69 (-13.70)	-1.66 (-9.19)	-1.77 (-10.55)
	Travel time	-2.30 (-34.37)	-3.10 (-17.36)	-2.74 (-16.61)	-2.41 (-38.32)	-2.95 (-18.39)	-2.75 (-18.85)
		Income Groups					
Middle Income Group as reference level	Lower Middle Income x Travel time	-0.138 (-1.757)	0.052 (0.22)	-0.33 (-1.38)	-0.137 (-1.72)	-0.73 (-3.05)	-0.91 (-3.72)
	Upper Middle Income x Travel time	-0.197 (-2.99)	0.088 (0.486)	-0.120 (-0.76)	-0.153 (-2.39)	-0.566 (-3.29)	-0.53 (-3.45)
	rho square (null)	0.508	0.698	0.677	0.512	0.711	0.685

Table 10: Estimated Betas, Work (2018, 2022 ODiN), Full vs 45 Alt Choice set (Random and Stratified Importance Sampling))

Estimated Betas, Work (2018, 2022 ODiN), Full vs 45 Alt Choice set (RS and SIS)							
		2018			**2022**		
		Full	**45 RS**	**45 SIS**	**Full**	**45 RS**	**45 SIS**
Note	Name	Parameter Values (t-test values)			Parameter Values (t-test values)		
Male as reference level	Female Gender x Travel Time	-0.18 (-4.24)	-0.32 (-4.39)	-0.25 (-3.55)	-0.13 (-2.71)	-0.32 (-3.69)	-0.156 (-1.87)
Parameters for logarithmic values	Jobs	1.14 (45.4)	1.16 (42.4)	1.16 (42.66)	1.09 (36.3)	1.12 (34.1)	1.10 (34.27)
	SCAE	-0.68 (-6.79)	-0.52 (-5.48)	-0.67 (-7.09)	-0.71 (-7.14)	-0.68 (-6.29)	-0.767 (-7.15)
	Travel time	-2.01 (-20.6)	-2.21 (-11.5)	-2.35 (-12.1)	-1.99 (-19.2)	-2.60 (-12.2)	-2.317 (-12)
		Age Groups					
Age group 65+ as reference level	Age 25 - 44 x Travel time	0.57 (6.09)	0.74 (3.87)	0.68 (3.97)	0.46 (4.76)	0.90 (4.39)	0.69 (3.82)
	Age 45-65 x Travel time	0.38 (4.08)	0.55 (2.85)	0.48 (2.81)	0.30 (3.16)	0.73 (3.52)	0.45 (2.44)
		Income Groups					
Middle Income Group as reference level	Low Income x Travel time	0.323 (2.82)	0.10 (0.53)	0.428 (2.26)	0.299 (2.3)	0.49 (2.41)	0.39 (2.11)
	Upper Middle Income x Travel time	0.156 (2.62)	0.07 (0.67)	0.33 (2.98)	0.117 (1.74)	0.192 (1.60)	0.13 (1.07)
	High Income x Travel time	0.26 (3.66)	0.25 (2.01)	0.44 (3.46)	0.23 (2.73)	0.195 (1.24)	0.197 (1.37)
	rho square (null)	0.215	0.32	0.31	0.211	0.319	0.304

Table 11: Estimated Betas, Primary education (2018 ODiN), Full choice set

Estimated Betas, Primary education (2018 ODiN), Full choice set		
		2018
Note	**Name**	**Parameter value (robust t-test value)**
Parameters for logarithmic values	Primary enrolments	0.59 (7.97)
	SCAE	-0.22 (-1.23)
	Travel time	-1.85 (-20.81)
	Income Groups	
Middle Income group as reference level	Upper Middle Income x TT	-0.206 (-2.117)
rho square (null)		0.565

Table 12: Estimated Betas, Secondary+ (2018, 2022 ODiN), Full vs 45 Alt Choice set (Random and Stratified Importance Sampling)

Estimated Betas, Secondary+ (2018, 2022 ODiN), Full vs 45 Alt Choice set (RS and SIS)							
		2018			**2022**		
		Full	**45 RS**	**45 SIS**	**Full**	**45 RS**	**45 SIS**
Note	Name	Parameter Values (t-test values)			Parameter Values (t-test values)		
Parameters for logarithmic values	Secondary+ Enrollments	0.42 (7.14)	0.397 (6.6)	0.399 (6.822)	0.49 (6.08)	0.46 (5.6)	0.49 (6.28)
	SCAE	0.62 (4.33)	0.67 (4.55)	0.72 (4.61)	0.40 (2.07)	0.535 (2.73)	0.48 (2.51)
	Travel time	-1.70 (-13.8)	-1.74 (-8.44)	-1.97 (-7.77)	-1.78 (-11.7)	-1.94 (-8.71)	-1.89 (-7.13)
		Migration Background Groups					
Dutch as reference level	Non Western x Travel time	0.25 (2.11)	0.36 (2.07)	0.488 (2.68)	0.13 (0.79)	0.02 (0.08)	-0.12 (-0.46)
		Income Groups					
Middle Income Group as reference level	Low Income x Travel time	0.273 (1.62)	-0.01 (-0.03)	0.38 (1.31)	0.645 (3.35)	0.74 (2.73)	0.902 (2.94)
	rho square (null)	**0.227**	**0.277**	**0.273**	**0.222**	**0.28**	**0.276**