Delft University of Technology

Master Thesis

Modeling a Tour Frequency Prediction tool for a Tour-Based Transport Model in the Netherlands

Author: Geraldo Çyrbja

- July 4, 2025

University Supervisors:
Dr. Maaike Snelder
Dr. ir. Alexandra Gavriilidou

Company Supervisor:
Dr. ir. Erik de Romph

Preface

This thesis marks the final step of my Master's program in Civil Engineering at TU Delft, with a specialization in Traffic and Transport. It represents not only the peak of my academic journey, but also a valuable learning experience that has deepened my understanding of transport modeling and its significance in the context of sustainable mobility planning.

The idea for this research originated from my strong interest in how data and models can help us understand complex human behavior and support better decision-making in urban engineering. I am especially grateful to Dr. Ir. Erik de Romph for offering me his invaluable expertise in the field and the opportunity to pursue this research topic within Haskoning, where I could explore the development of tour-based demand models in an industry context. Working on tour-based modeling allowed me to combine technical skills that I had acquired during my studies with behavioral insights, which I found both challenging and interesting.

This research would not have been possible without the support and guidance of several individuals. I would like to express my sincere gratitude also to my daily supervisor, Dr. Ir. Alexandra Gavriilidou, whose continuous support, availability, and valuable expertise in discrete choice modeling were crucial throughout the entire thesis process. Her detailed feedback and suggestions significantly strengthened both the methodology and the report of this research. I would also like to thank the chair of my committee, Dr. Maaike Snelder, for her guidance and insights, which helped me maintain a clear direction in the research. I am also grateful to Haskoning for giving me the opportunity to carry out this research within their prestigious organization and for offering me the resources that made this work possible.

Last but not least, I want to thank my family for their unwavering support during this intense period of research and throughout my whole academic journey. Their belief in me, even during the most uncertain moments, gave me the strength and motivation to keep going. None of my achievements would have been possible without their support, and for that, I am deeply grateful. I also want to thank my friends for being by my side, bringing laughter and unforgettable moments that made even the hardest times more beautiful.

I hope that the findings of this research contribute to the ongoing development of more behaviorally sound transport models and offer a small step forward in the path toward more sustainable mobility solutions.

Yours sincerely,

Geraldo Çyrbja

Summary

Understanding and forecasting travel behavior is essential for sustainable transport planning, especially due to the fact that urban development is increasing and mobility needs are increasing with it. Traditional trip-based models, which model trips as isolated events, fail to capture the complexity and constraints of real-life travel behavior. As a result, there has been a shift toward more behaviorally realistic models, such as tour-based and activity-based models, which model travel demand based on linked trip chains and daily activity schedules. Tour-based models stand in between the aggregated simplicity of trip-based models and the disaggregated complexity of agent-based models. By modeling tours as sequences of trips that begin and end at home (or work), they reflect spatial and temporal constraints better. One of the main components of a tour-based model is the tour generation based on explanatory variables (e.g., age, occupation, income) for the population input characterized by those attributes. The main motive behind travel is activity participation, so constructing the tour patterns for different individuals requires a thorough understanding of the factors that affect the participation of these individuals in different activities.

This research focuses on developing a tour generation tool by providing a robust model structure and a set of explanatory variables that generate the daily tours taken by specific population segments, mimicking real-world behavior. The model is developed and validated using the Dutch ODiN (Onderweg in Nederland) travel survey, aiming to better represent behavioral diversity between population segments. Given the complex transport behavior in the Netherlands, supported by an advanced multimodal transport network, a robust tour generation model can serve as a valuable tool for making informed decisions and guide policy-making in the field by giving insights on what affects the generation of travel for different individuals. The study investigates which demographic, socio-economic, and spatial attributes influence tour generation for enhanced predictive capabilities. Furthermore, the working-from-home phenomenon is also explored, as it is becoming an increasing trend after the COVID-19 pandemic.

The literature review provided a couple of modeling approaches commonly used for the generation of tours from explanatory variables. These were mainly divided into two categories: regression-based and discrete choice models. After reviewing these methods, a Discrete Choice Modeling (DCM) approach was chosen for this research due to its strong behavioral foundation and ability to represent individual-level decision-making processes. DCMs provide a clear and interpretable understanding of how different factors influence the choices made by travelers. To develop the model, a choice structure with alternatives that the decision-maker can choose and a set of explanatory attributes are required.

The ODiN data contains full daily travel records of the surveyed individuals, which are used to form activity trip chains (tours) that start and end at home (e.g., Home-Work-Shop-Home) as a reference location where individuals depart from and return to after carrying out certain activities. The number of different trip-chains that appear in the data is very large due to the combinatorial complexity. Therefore, based on the literature (Yagi and Mohammadian, 2008), a categorization of the tours is carried out by assigning a primary activity for each of them based on a priority order (Work-Education-Shop-Leisure-Escort-Other), assuming that the highest ranking activity in each trip chain determines the nature of the tour, and the other activities are considered secondary. This categorization is used to separate the choice structure into two stages to reduce the number of alternatives and distinguish between the effects of different attributes in different tour types. In the first stage, the choice is made for a daily pattern between staying at home and choosing the

primary purposes for which a tour will be made, see Fig. S1. After that, in the second stage a separate choice is modeled for each of the 6 primary purposes to choose the trip-chain (or multiple trip chains) for the primary purposes that were chosen in the first stage. So, the final result of the combined model for each individual characterized by certain attributes, is the full set of trip chains that is modeled with the probability of choosing each of them. The set of explanatory attributes that are used to form the utility functions of each alternative are: age, gender, occupation, household income and size, car ownership, urban level (based on household density), and zone level (public transport stops and population) of the place of residence. An alternative choice structure of nested logit type is defined for each model to compare with the base multinomial structure to provide conclusions on the best model structure. The nests for the first-stage model are home (stay at home and work-from-home) and travel (all alternatives with tours). In the second stage for each primary purpose, the 1-trip chain and 2+ trip chain alternatives are grouped in two different nests.

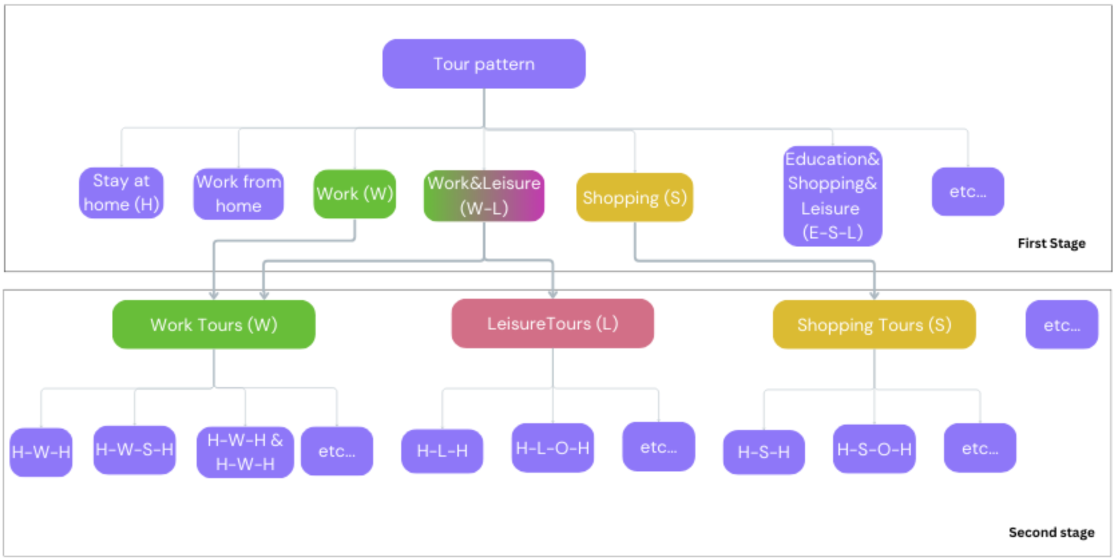


Figure S1: Choice structure with alternative examples (Multinomial Logit)

The models are trained on data from 2022-2023 as the travel behavior after the pandemic changed significantly, ensuring the validity of the model for the current and near-future behavior. The results revealed important insights on the influence of different attributes in the generation of tours and demonstrated high predictive accuracy when used to model the choices made in a validation dataset that was not used for training. The first-stage model results showed that all attributes, except the zone level, have a significant effect on choosing a daily pattern. The nested model outperformed the multinomial logit for the first-stage model in terms of predictive capabilities. The second-stage models results indicate that different attributes have an effect when choosing the trip chains for different primary purposes, with age and occupation appearing most consistently. The multinomial logit structure is chosen for all second-stage models as the nested models did not improve over any of them.

The combined model is tested in the validation dataset to predict the choices that are then compared to the observed trip-chain counts. The predictions were highly accurate at an aggregate level, but significant over- or underestimations were revealed when looking at certain subgroups (students, elderly). A respecification of the first-stage model with alternative-specific parameters for 7 alternatives to distinguish the effect of attributes across different alternatives significantly improved the performance of the model within subgroups, while keeping the overall aggregated model performance high. Therefore, the final first-stage model is this respecified model with alternative-specific parameters for 7 alternatives with an MNL structure (the nest coefficient for the nested structure is not significantly different from 1), see Fig. S2 (green means positive and red negative effect on utilities).

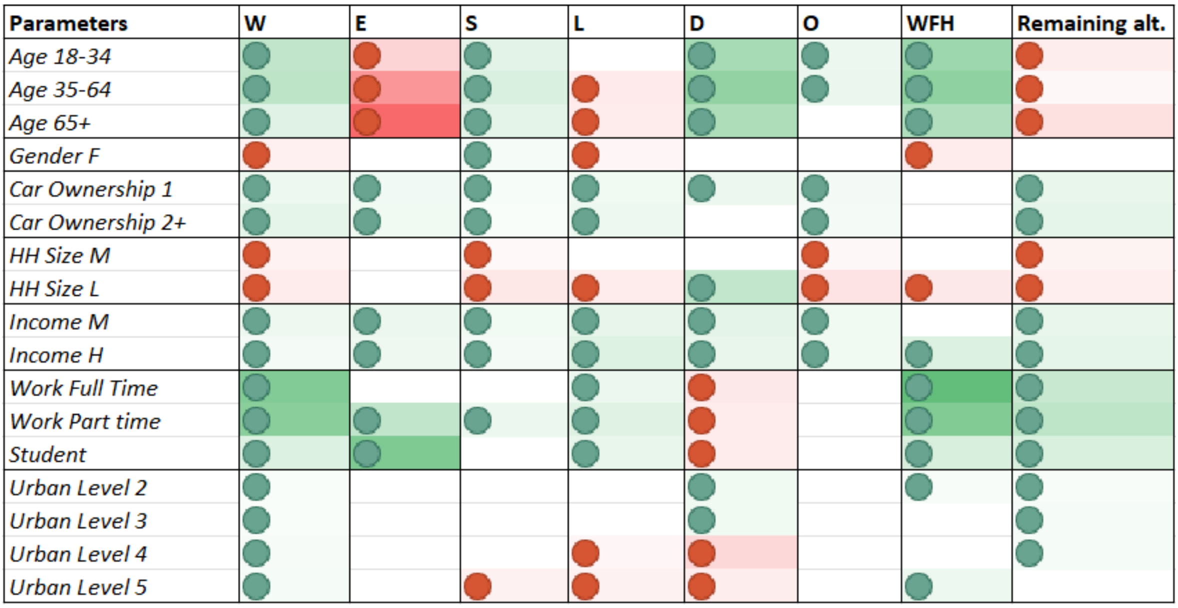


Figure S2: The effect of parameters in the alternatives of the first stage model with alternative-specific parameters

A case study compared the performance of the developed discrete choice model with a current frequency-based method used in practice (de Romph, 2021) for estimating the frequencies of different trip chains for an input population and comparing them with the observed counts in a validation dataset not used for training any of the models. The discrete choice model consistently estimated the observed tour counts quite well, while the frequency-based method tended to overpredict the less frequent trip chains, see Fig. S3.

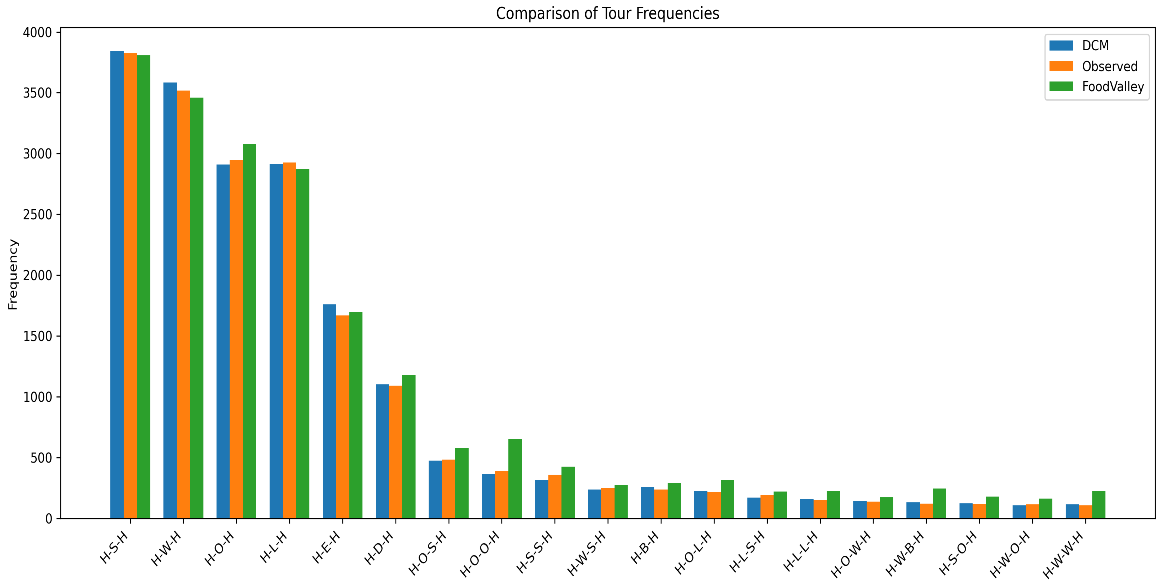


Figure S3: Observed and estimated counts of trip chains modeled from both methods

In conclusion, the discrete choice model with a two-stage choice structure of multinomial logit type provides a robust framework to explicitly model the trip chains that are generated by individuals with different characteristics. All the selected explanatory attributes revealed significant effects on the choices made on different parts of the model, except for the zone level attribute, suggesting that public transit accessibility does not have a strong influence on tour generation. The validity of the model across different regions or future time frames with varying travel behavior can be ensured by re-estimating the model with data relevant to the desired context. Future research regarding this topic can explore different approaches for modeling trip chains, instead of explicitly including them as alternatives. Furthermore, the exploration of additional choice structures or the inclusion of work-based tours as a separate category could further improve the explanatory power of these models.

Contents

List of Figures

List of Tables

1 Introduction

The prosperity and development of society throughout history have always depended on the effectiveness of its transportation systems. Population growth and high urbanization rates, especially in developed countries, increase the reliance on transportation systems, and transport policymakers require deep insights into travel behavior to offer efficient solutions that ensure sustainable development. Transport models are widely used instruments by transport experts to make informed decisions about policies and investments in this important field (PTV Group, 2024). Travel demand is shaped by the spatial and temporal distribution of the activities that individuals participate in, and for that reason, it is a derived demand from the need to reach activity locations (Heggie and Jones, 1978). The conventional 4-step trip-based approaches were developed to predict travel demand with focus on providing the necessary transportation supply to support it. They model trips as isolated events at an aggregate-level and do not take into consideration spatial and temporal constraints, presenting several limitations with respect to travel behavior. Trip-based models also present difficulties in modeling future situations that lack evidence from practice in the present due to their limited behavioral realism (Dumbliauskas, 2019).

Transport planners have lately been shifting their focus from supply-oriented to demand-oriented planning as continuously increasing the capacity of infrastructure is not a sustainable solution with the rapidly growing mobility needs (Ipek Sener et al., 2009). Therefore, there is a growing interest in more advanced disaggregate-level methods, such as tour-based modeling, that have tours as their unit of travel and model travel demand driven by the activity participation principle. Tours combine single trips in a trip chain that start and end at a specific location (usually home). This family of models tackles the limitation of trip-based models that treat individual trips as independent decisions and the effects of other activity decisions not being considered, by offering a more realistic framework that imposes spatial and temporal constraints in trip chains (Rossi and Shiftan, 1997). However, most tour-based models still treat tours as independent, neglecting the interdependencies beyond tours (Bowman and Ben-Akiva, 2001). They follow a more disaggregate approach than trip-based models (aggregated), usually macroscopic disaggregation at the level of population segments (e.g., students, full-time workers, etc.) to simulate travel demand, allowing for behavioral differentiation without the complexity of fully microscopic simulation. Tour-based models present an advancement toward activity- and agent-based models characterized by further disaggregation. Activity-based models simulate the demand at the individual or household level by constructing their daily activity schedules and imposing time constraints (Dumbliauskas, 2019). Agent-based models simulate microscopic travel demand at individual level making sure that all tours and trips are feasible in time and space (Vovsha, 2019).

Constructing daily activity patterns of population segments by generating the activities and travel demand (tours) needed to reach those locations is a crucial part of tour-based models. This is usually carried out by a tour frequency module, in which different types of tours are generated by means of explanatory variables (household information, land use, mobility). While tour-based models are being actively explored, many limitations still arise in identifying and modeling a large range of factors that influence the demand for different activities. These tour generation models are usually constrained by limited demographic variables and overlook dynamic factors and accessibility effects. The aim of this research is to develop and evaluate a model that predicts the frequencies of different types of tours across population segments. By evaluating its performance, the study aims to assess the effectiveness of the methodological approach and model structure for supporting travel demand generation for a tour-based transport model.

The research focuses on the Dutch context. The Netherlands has a well-developed transportation network with a diverse set of modes that provide many travel options, but at the same time, it faces many challenges, especially in increasing the efficiency and reliability of multi-modal travel. A reliable tour-based model that replicates and analyzes complex behavior can be a very useful tool to investigate the critical points and guide the development of transport solutions. However, the development of this model requires a reliable tour generation tool to model the travel demand as close to reality as possible. People often combine several trips throughout the day (e.g., commuting to work-grocery shopping-gym-home) and that affects mode, time choice, and travel frequency. There are a large number of trip purpose combinations that can make up a tour and people have different schedules and habits, making the demand modeling process quite complex.

The focus of this research is the enhancement of the explanatory power of a tour generation model by exploring explanatory attributes and model structures. By analyzing the factors that influence how often individuals undertake different types of tours, this study aims to develop a model that captures behavioral variations across population segments. The model development and validation will be supported by travel data sources collected in the Netherlands, primarily from the ODiN (Onderweg in Nederland) travel survey conducted by the Central Bureau of Statistics (CBS) (Centraal Bureau voor de Statistiek, 2024). Key demographic, socio-economic, and accessibility-related attributes will be examined to model their effects on tour generation. The research questions to achieve the aim of this research are described in section 1.1 and the contributions to the society in section 1.2.

1.1 Research questions

The development and comparison of different model structures for tour generation and testing the explanatory power of different attributes form the basis to achieve the aim of the study and answer the main and sub-research question:

- **Which model structure and predictors should be used to predict the frequencies of various tour types across population segments for tour generation?**

  1. Which rules should be used to form tours from trip data and aggregate different types of trip chains (travel purpose combinations)?

     Subquestion 1 addresses the formation of tours from trip data and categorization of trip chains (tours) by aggregating them into representative categories for analyzing and modeling them. A framework will be developed in Chapter 3, supported by a literature review in Chapter 2.

  2. Which methodological approach is most suitable for estimating the frequencies of different tour types taken by specific population segments?

     Subquestion 2 focuses on identifying the most suitable modeling approach for estimating tour frequencies across segments (e.g. regression-based or machine-learning methods), which will be used to develop the models that will be assessed. Literature review (Chapter 2) will be the foundation for exploring and choosing a suitable modeling approach and possible modeling structures.

  3. Which explanatory attributes available in the data can be used to supply the model development?

     Subquestion 3 explores explanatory variables that the available data can provide to explain travel generation behaviour. The output of this subquestion will serve as input for the set of explanatory variables that are used and tested in the models. It is supported by the literature review and mainly from the data analysis chapter (Chapter 4).

4. Which performance metrics should be utilized to measure the models' abilities to predict tour frequencies?

   Subquestion 4 is critical in answering the main question by providing tools to measure the performance of the developed models. By defining evaluation metrics and validation techniques, this step determines how well the model predictions align with observed travel patterns and is crucial in answering the main research question. Performance indicators are supported by literature and defined in the methodology (Chapter 3).

1.2 Scientific and societal contribution

Significant advancements have been achieved in existing literature (see Chapter 2) about tour-based demand models. However, a critical gap seems to remain in defining a robust framework and a set of explanatory variables that accurately predict how often people perform certain types of tours. Most of the state of art consists of models that predict tours primarily aimed at the primary activity of a tour and limited in terms of secondary activities, often neglecting the need for estimating the frequency of different trip chains by making a distinction in the order of activities. This research aims to bridge this gap by increasing the predictive and explanatory power of tour generation models by developing a tool that predicts the frequency of various tour types (trip chains) across distinct population segments, moving beyond binary decisions or simplified explanatory variables. Furthermore, the phenomenon of working from home will be explicitly modeled as it is becoming more and more common and could lead to significant effects on traffic demand.

The outcome of this thesis is a crucial component for transport models, which are widely used instruments for evaluating the performance of transportation systems and aiding the decision-making process in this field. Transport is essential in connecting people, goods, and services, making social and economic development highly dependent on it. Therefore, instruments and frameworks that support the development in this field affect a large range of entities.

The research community is interested in advancing knowledge in transportation modeling and travel behavior. Providing an innovative framework that increases the predictive power of travel demand models and analyzing factors that drive travel behavior strengthens the credibility of transport models. The outcomes of this research or a transport model that might implement this tour generation model might be used by policymakers in the future to analyze transport demand and make decisions based on the model insights. Society can be impacted by this study if the transport model that will be developed is used in designing and operating transportation infrastructure and services. The quality of transport and accessibility has a direct effect on the quality of life of the citizens. Therefore, maintaining high research standards and integrity is of great importance, considering the wide impacts that this study can have on society.

1.3 Thesis Outline

This thesis is structured into five chapters, each contributing to the development, application, and evaluation of a tour generation model. Chapter 2 covers literature review on tour-based models, existing approaches for estimating tours, and lays out the theoretical foundation for building the methodology. The methodology chapter 3 follows by defining the modeling framework and performance indicators, setting the path to answer the main and sub-research questions. Chapter 4 describes the data analysis and preparation process for model estimation. Chapter 5 presents the results of model output, including performance analysis and validation. Furthermore, a case study demonstrating the application of the model in a practical context is carried out at the end of this chapter to further assess the model relevance and applicability by comparing it to an existing approach. Finally, chapter 6 provides the answers to the research questions, discusses the findings, and lays out recommendations for future research and possible model improvements.

2 Literature review

This chapter discusses existing literature on different demand modeling approaches for tour-based models. The structure of a tour-based model within the context of this research and its components are discussed in Section 2.1. This review explores tour classification methods used in practice to lower the complexity of trip chain combinations for efficient model development (Section 2.2). Furthermore, current frameworks used for predicting tour frequencies are reviewed (Section 2.3) and analyzed to identify current practices and gaps in the methodologies. The literature provides the foundation for identifying the relevance of this research and supports the methodology that is used to achieve the objective.

2.1 Tour-based model

Tour-based travel demand models use tours as their unit of travel, being a sequence of trips that begin and end at the same location, usually home or work (Ipek Sener et al., 2009). The primary purpose of the tour, usually being the longer-lasting activity, defines the type of the tour (e.g. work or shopping tour), and the other activities within a tour are considered secondary. Tour-based models follow a macroscopic disaggregate approach usually at the population segment level, to simulate travel demand.

A tour-based model that can integrate the tour generation tool consists of several connected modules that model travel demand, see Fig. 2.1 (de Romph, 2021). The population of the study area and its characteristics are a crucial stepping stone to simulate travel demand in tour-based models. The population synthesizer generates a synthetic population for the study area based on demographic and socio-economic data for each zone, whose individuals contain several attributes (e.g. age, occupation, car ownership) that influence their daily travel, and which are used for demand generation in later stages. The synthetic population is used as input for the tour generation module to simulate the daily tours performed by different population segments that individuals are part of, which is one of the core building blocks of these models and has a detrimental effect on the quality of the output. The tour frequency model predicts how often individuals of a certain population segment undertake certain trip chains based on some explanatory attributes. Different tour types are generated by different circumstances and therefore separate models per tour type are usually used to estimate the frequency that an individual from a specific population segment, characterized by specific attributes, takes a tour of a certain type. It is often modeled from household data, and more advanced prototypes try to incorporate land-use and accessibility effects in the generation of travel (Ida Kristoffersson and Algers, 2020). After the generation of tours, the destination choice is carried out for the primary activity of each tour, followed by the tour-level mode choice based on the destination of the primary activity. The destinations of the secondary activities are chosen in the next step and the time of the day choice for each part (trip) of the tour is carried out in the last step (de Romph, 2021).

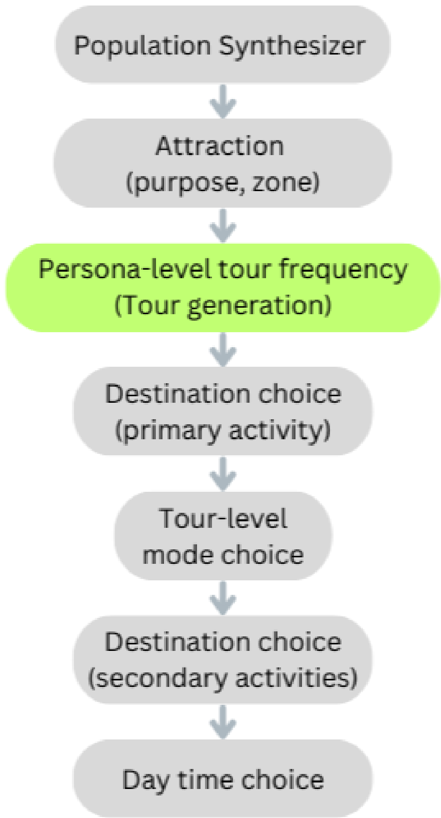


Figure 2.1: Haskoning's Tour model structure (de Romph, 2021)

2.2 Tour classification

The aggregation of tours into categories is necessary to tackle the combinatorial complexity of trip chains that can be formed of many combinations of trip purposes that travelers take. Primerano et al. (2008) established a link between primary and secondary activities and explored the share of different trip-chain types across distinct population segments. Yagi and Mohammadian (2008) defined four main primary purposes with priorities to rank their importance in a tour in this order: work, school, maintenance, and discretionary. Bowman and Ben-Akiva (2001) defined 9 tour types by purpose, number, and sequence of stops, 5 being different configurations (presence of intermediate stops) of work-related tours, 2 school-related tours, and two categories (one with and one without an intermediate stop) of tours with main activity other than school and work. The SACSIM tour-based model (Ipek Sener et al., 2009) defined 7 main tour purposes (work, school, personal business, escort, shopping, meal, social interaction) and a set of 2080 possible tours with combinations of these purposes (characterized by the main purpose and intermediate stops).

Nowrouzian and Srinivasan (2012) categorized tours on four criteria: purpose representing the main activity, number of stops (complexity), flexibility, and travel party composition (solo or joint travel). Pirra and Diana (2016) carried out a tour classification for the U.S. using clustering techniques and national travel survey data. The clusters that were defined considering activity types, travel modes, and socio-demographic characteristics are: tours with long duration work activities, short tours (usually secondary tours within the day), tours made by elders or retired people, and tours performed by young people (usually school tours).

The concept of a primary activity for tours with multiple stops to categorize them in meaningful categories is widely supported in literature and is therefore adopted and used in Chapter 3.

2.3 Existing tour frequency models

Literature documents different modeling approaches for the tour generation component of the model. The vast majority of studies employ regression-based (Subsection 2.3.1) or discrete choice (Subsection 2.3.2) methods, but innovative machine learning approaches have also been gaining interest from researchers, even though literature on this family of models is quite limited.

2.3.1 Tour-based regression demand models

Regression models are widely used in literature to simulate tour generation by fitting a model using explanatory attributes and travel survey data. Golob (2000) combined the trip generation and time use to capture trip chaining (tours) and activity participation effects for demand modeling. The methodology involved using endogenous (activity duration, work/nonwork tours, travel time, etc.) and exogenous (car ownership, income, etc.) variables to build a structural equation model and estimate it using the Maximum Likelihood method. It was one of the early alternatives that included trip chaining and time constraints in demand forecasting, rather than treating trips as isolated events. Krizek (2003) assessed the effect of land use mix and accessibility in the residential neighborhood in the tour formation and frequency of certain tour types by using regression models. The results showed that the neighborhood accessibility (easiness of reaching certain activities from a residential location) was a significant variable for estimating the frequency of simple commuting or maintenance (shopping, personal care, etc.) tours only.

Guzman et al. (2017) developed and validated a strategic travel demand model (LUTI model) for the capital of Colombia, Bogota. The model involved tour generation and attraction by using regression models fed with data from a cross-sectional mobility survey. Spatial, socio-economic factors and travel time budget were used as variables to estimate the amount of tours, which were simply divided into two categories (commuting and non-commuting tours). The model included several feedback loops, considering congestion effects, travel time, and costs. On the other hand, Dumbliauskas (2019) followed an empirical methodology by analyzing mobility survey data conducted in the city of Vilnius. This framework involves calculating probabilities of many different trip sequences (tours) that were present in the collected data, being taken by different population segments (by age and employment). These probabilities were then multiplied by the total number of residents that fall in each population segment to get the total travel demand for each tour type. A similar approach is followed by Hamad and Obaid (2022) by defining a mobility rate for a certain trip chain performed by a certain population segment and multiplying it with the size of that population in a specific zone to calculate the total production of that tour type.

Machine learning techniques have lately gained popularity in transportation modeling as data-driven approaches with high predictive power and the potential to capture complex relations between variables. Přibyl and Goulias (2005) utilized a decision tree algorithm (machine learning method) to simulate the activity patterns of different individuals throughout the day on a disaggregate level. Goel and Sinha (2008) developed a model that forecasts trip generation using an artificial neural network architecture. Two separate neural networks were developed on the production and attraction sides. Different input variables, such as population, number of cars, etc, made up the input layer of the network that is used to estimate the model and later predict the generation of tours. The results demonstrated high predictive power for this model due to its ability to capture non-linear relationships.

2.3.2 Discrete choice tour frequency models

Discrete choice modeling has been lately in the mainstream of travel demand modeling. These types of models make use of utility maximization theory to simulate the choice process of individuals by evaluating different alternatives based on relevant explanatory attributes. Festa et al. (2006) modeled the decision-making process of travelers (surveyed in a medium-sized urban area in South Italy) using Binomial logit models for 4 tour types (shopping, recreation, escort, maintenance) to decide if a person does that type of tour during a day or not (0/1+). Strong correlations were revealed between explanatory variables and travel demand behavior, but some variables were not significant for all types of tours. Yagi and Mohammadian (2008) proposed a two-tier nested logit model with the choice to go out or not in the upper nest and choosing between daily activity-pattern alternatives (defined by primary activity, tour type, and type and number of secondary tours) in the lower nest. Each pattern included a primary tour and a choice between 0/1/2+ secondary tours (maintenance, discretionary). The total choice set with different combinations of primary and secondary tours resulted in 121 activity pattern alternatives.

Esmael et al. (2009) also developed a hierarchical discrete choice model for tour generation in a nested logit form starting from the daily travel pattern (0, 1, 2 tours, tour with subtour) to tour type choice: primary and secondary tour type choosing from the available 9 alternatives characterized by three main tour purposes (work & school, maintenance, discretionary) and three intermediate stop scenarios (0/1/2+). Shams et al. (2018) used binomial logit models to estimate the daily tour rate (1 or 2+ tours) for commute and shopping tours only. Ida Kristoffersson and Algers (2020) introduced an upper limit to the number of tours that people make per day in the tour generation model of nested logit type. In the upper level, the number of tours per day is determined (0/1/2/3/4), and on the lower level, a tour pattern is chosen within the alternatives of tour categories. The number of tour pattern alternatives generated was 86 (1 stay-at-home, 10 one-tour, 50 two-tour, 24 three-tour and 1 four-tour patterns). It also included accessibility using a log-sum variable from lower nest destination and mode choices as a factor, which came out important for non-mandatory tours, but not significant for mandatory tours (work, school).

A prototype model for Boston developed by Bowman and Ben-Akiva (2001) carries out the tour generation process through a nested logit model, whose upper nest decides if the person stays home or travels and the lower nest under the travel option provides a set of 54 travel pattern alternatives to choose from for the whole day. The activity patterns are characterized by the primary activity of the day (home, work, school), primary tour type (9 alternatives explained in Section 2.2), and number and purpose of secondary tours (6 alternatives). The combinations of these attributes resulted in 54 daily activity patterns that generate the disaggregate travel demand.

San Francisco Transportation Authority (SFCTA) developed a tour-based model (SF-CHAMP) that contains a full-day tour pattern choice submodel of Nested Logit type (Ipek Sener et al., 2009). It predicts several dimensions such as the purpose and trip-chain type of the main tour and frequencies of home-based secondary tours (0, 1, 2+) and work-based sub tours (0, 1+). For intermediate stops (secondary activities) it considers the accessibility to retail and service locations within 15 minutes from home or work. The data used included socio-economic, land-use (connectivity, parking, school, area type), and transportation service level (accessibility, transit times).

The SACSIM tour-based model designed for the region of Sacramento in California employs an activity simulator (DaySim) that generates daily travel (tours and trips within tours) for each individual in the population (Ipek Sener et al., 2009). The pattern generation model predicts the number of tours that a person takes from seven purposes (work, school, personal business, escort, shopping, meal, social interaction) and intermediate stops from the same purposes, resulting in a set of 2080 feasible alternatives of combinations of 0/1+ tours and 0/1+ stops. The choice model is of multinomial logit type and the base alternative is staying at home. The output of this model is used from a subsequent model to determine the exact number of tours (1, 2 or 3) from the seven purposes for which a tour was predicted in the previous model. Accessibility to service locations and residence neighborhood density are considered in shopping tours or intermediate stops.

The journey frequency module of the NYBPM (New York Best Practice Model) tour-based model considers three person types (worker, non-worker, child) and six purposes (to work, to school, to university, maintenance, discretionary, at work as a non-home based tour), resulting in a total of thirteen tour-frequency sub-models tailored to the type of person and the main tour purpose. The sub-models related to school, university, and discretionary activities are of binary logit type by choosing between 0 or 1+ tours of that type. The journeys to work and maintenance activities are predicted using multinomial logit models to choose between 0/1/2+ tours to work and 0/1/2/3+ tours for maintenance (Ipek Sener et al., 2009).

Dutch national transport model LMS also includes a discrete choice model to calculate tour frequencies for 12 travel purposes that are used to construct the travel patterns of the population. For each purpose, there is a two-step structure, the first being a choice between 0 and 1+ tours and a subsequent stop/repeat sub model that determines if exactly one or more tours of that type are made and so on (de Jong et al., 2007). The utility functions are based on personal and household characteristics (household composition, car ownership, age, gender, driving license, education, income, and degree of urbanization). There are two utility functions defined, one for the 0/1+ sub-model and one for the stop/repeat sub-model. The coefficients are estimated for the 0 and stop alternatives, while utilities of 1+ and repeat alternatives are set to 0. Similarly, a transport model of the city of London estimates the frequencies of different tour purposes with separate models that contain two sub-models each (Patruni et al., 2021). The first sub-model decides if no tours or at least one tour (0/1+) of a given purpose is made on an average weekday and the second one predicts how many of those tours are made.

2.3.3 Modeling approach for this research

This research adopts a Discrete Choice Modeling (DCM) approach for predicting tour frequencies, even though other methods were reviewed. DCMs are rooted in behavioral theory and are suited for representing individual-level decision-making processes. They allow for probabilistic modeling of choices among discrete alternatives, such as different types of daily tour patterns based on utility maximization principles.

One key advantage of DCMs over regression-based methods is their ability to model the decisions more explicitly, incorporating factors like relative utility, choice set composition, and nested relationships between alternatives. Unlike machine learning techniques, which often operate as black-box models, DCMs provide interpretable parameters, making it possible to understand the impact of socio-demographic attributes, spatial and accessibility indicators on tour generation. In transport planning, understanding the reasons behind travel behavior is just as important as forecasting outcomes. Therefore, DCMs seem to offer an effective balance between predictive capability and behavioral insight, making them a suitable choice for this study.

2.4 Explanatory attributes

The explanatory attributes are used in the models to predict the daily tour patterns of population segments. These attributes can be related to the socio-demographics of the population (household composition, income, employment, age, gender, driving license availability etc.), land-use (urban form), or accessibility (car availability, transit accessibility, etc.). They capture the activity participation of the population at a disaggregate level to construct their travel patterns. The reviewed models usually utilize different attributes to predict different types of tours. However, the main attributes present in literature and the studies or models in which they appear are summarized in Table 2.1. As can be seen, household composition (number of adults and children), occupation (employment, student), gender, age and car availability appear in most of the studies as significant attributes. The income variable is also very common, sometimes appearing as household income and sometimes as personal income. The urban form or density of the residential location has been included in a few models, proving to be a significant attribute in tour frequency prediction. Accessibility logsums are also found in the utility functions of tour frequencies in a nested logit models to represent the accessibility in terms of the effect of lower-level mode and destination alternatives on higher-level tour pattern choices.

Table 2.1: Explanatory attributes used in existing models

	HH comp.	**HH inc.**	**Emplo-yment**	**Stu-dent**	**Inc-ome**	**Gen-der**	**Age**	**Lice-nce**	**Car**	**Residen-ce urban level**
Yagi and Mohammadian (2008)	x	x	x	x		x	x		x	x
Festa et al. (2006)	x		x			x	x	x	x	
Ida Kristoffersson and Algers (2020)	x		x		x	x	x	x	x	
Esmael et al. (2009)	x		x	x		x	x			
Bowman and Ben-Akiva (2001)	x		x	x	x	x	x			
SFCHAMP - Ipek Sener et al. (2009)	x	x	x	x		x	x		x	x
SACSIM - Ipek Sener et al. (2009)	x	x	x	x		x	x		x	x
NYBPM - Ipek Sener et al. (2009)	x	x	x	x			x		x	x
LMS - Rijkswaterstaat (2021)	x		x	x	x	x	x	x	x	x

2.5 Calibration & Validation

The models that are developed require calibration and validation to ensure the reliability and quality of the predictions. The tour frequency models are trained on datasets collected from travel surveys. Therefore, the model parameters can be calibrated based on the travel choices made by the survey respondents. The parameters are estimated using the Maximum Likelihood method that maximizes the probability predict the choices made by the participants of the survey (Festa et al., 2006). The likelihood function is represented by the product of the probabilities that the model predicts in choosing the right alternatives that are observed from the dataset, and the model parameters are optimized to maximize this function. Additionally, formal and informal tests are applied to assess the estimated parameters (Festa et al., 2006). Formal tests refer to the statistical significance of the parameters as a measure of effectivity of the explanatory power of attributes on the predictions. Informal tests on coefficients allow to check if they align with theoretical expectations in decreasing or increasing the utility.

The model predictions are usually validated by comparing them to the observed travel patterns from the collected data (Ipek Sener et al., 2009) as a measure of accuracy. Internal validation is carried out by splitting the dataset from the surveys into a training and validation subset to assess the performance of the model in a similar setting (Groeneveld et al., 2023). On the other hand, external validation is necessary to test the model on external datasets with a different setting to evaluate how the model performs on unseen data and provide a measure of the transferability of the model.

2.6 Conclusion literature

This chapter reviewed key concepts and methodologies from existing literature that form the foundation for this research. First, an overview of the components of tour-based models was provided to clarify how the developed model fits into the bigger picture and connects with other modules. Various approaches for classifying trip chains were examined, with this study adopting the concept of activity prioritization as the basis for tour formation and classification, (described in Section 3.2). Two broad categories of modeling approaches, regression-based and discrete choice models were discussed. While regression-based methods offer simplicity, discrete choice models are preferred in this study due to their stronger behavioral foundation and flexibility in handling choice structures (see section 2.3.3 for more details).

A set of explanatory attributes from previous studies helped identify commonly used socio-demographic and accessibility factors influencing tour behavior. These will form the starting point for the set of attributes in this model, which will be refined and expanded based on the structure and availability of the data in the ODiN dataset. Finally, the review of validation techniques forms the basis for the model performance evaluation methodology. This literature review forms the theoretical and empirical foundation necessary to proceed with model development in the following chapters.

3 Research methodology

This chapter describes the data sources and methodology that is followed to construct model prototypes that will estimate the tour frequencies to generate the travel demand of a typical working day (weekday) for the tour-based model. The methodology will provide a framework to answer the research questions and choose the best model structure, see the flowchart in Fig. 3.1 for an overview of the methodology. The data and tour definition are discussed in sections 3.1 and 3.2 respectively. Furthermore, the model specification is explained in section 3.3, where important modeling choices and structures are described in detail, while section 3.4 outlines the performance evaluation and a case study.

3.1 Data

The main dataset used in this research is the Onderweg in Nederland (ODiN), the official Dutch national travel survey conducted annually by the Centraal Bureau voor de Statistiek (Centraal Bureau voor de Statistiek, 2024). ODiN provides detailed insights into the travel behavior of the Dutch population and serves as a key input for transport planning and modeling studies in the Netherlands. It is a cross-sectional household travel survey in which respondents are asked to report all trips made on a specific day. The dataset includes both individual-level and household-level information, such as:

- Socio-demographics: age, gender, education, income, household composition, employment status.
- Mobility characteristics: car and bicycle ownership, driver's license possession, access to public transport.
- Trip details: purpose, mode, duration, start and end time, origin and destination zones.
- Population segment factor for the surveyed person: indicates the weight of that person in the population as the sample is representative of the whole population of the Netherlands.

Each reported trip is associated with a specific travel purpose, allowing the reconstruction of tour patterns for each individual. For this research, data from multiple years (2018–2023) is available for use. The data will be processed and analyzed to identify trends and suitability for modeling. The richness and consistency of ODiN make it suitable for disaggregate-level tour frequency modeling.

3.2 Tour definition and classification

The formation and classification of tours (yellow box in Fig. 3.1) from the survey data is the first step of the modeling phase. Tours are defined as a sequence of trips that start and end at a specific location, usually home. There is a large number of tour types that can be constructed by combining different activities that people perform during the day in trip chains. Therefore, a framework is needed to define the types of tours that will be modeled and what characterizes them. The ODiN data consists of all the trips that an individual undertakes during a day and the reasons for making those trips. For each trip reported from the data, a purpose is given based on the reason for the trip (Work (W), Business (B), Education (E), Shopping (S), Leisure (L), Drop-off (D), Other (O), Home (H)). These purposes summarize most of the trip reasons in these meaningful categories and still keep the trip chain complexity acceptable as adding more detailed purposes results in additional combinations of trip chains and a less efficient model.

Tours are formed as Home-based (HB) tours that start and end at home, because that is usually the location where people initiate movements and return to at the end of a tour and the end of the day. Tours with only one activity are straightforward to classify based on the purpose of their only destination, but for tours with two or more visited activities, the primary purpose for the tour should be determined. This choice assumes that in case individuals engage in multiple activities within a tour, there is a primary activity which has a higher importance (e.g. work) and restricts the other activities, so the rest of them come as secondary with respect to the primary one. This key activity is used to classify tours into categories with similar behavior. Therefore, tours are given a primary purpose based on a priority hierarchy of this order: work, business, education, shopping, leisure, escort, and other (Patruni et al., 2021). Apart from the primary tour purpose, the number and purpose of intermediate stops are important attributes for defining the tour types. In tours with two or more visited destinations, the activities other than the main activity are considered as secondary stops as they belong to a lower rank in the hierarchy of the purpose importance. The intermediate stops can be made before or after the main activity and those are considered as distinct types of tours (Home-Shop-Work-Home is different from Home-Work-Shop-Home). The formed and categorized trip-chains are the output of this stage and are used for analysis and model input into the next steps (see Fig. 3.1).

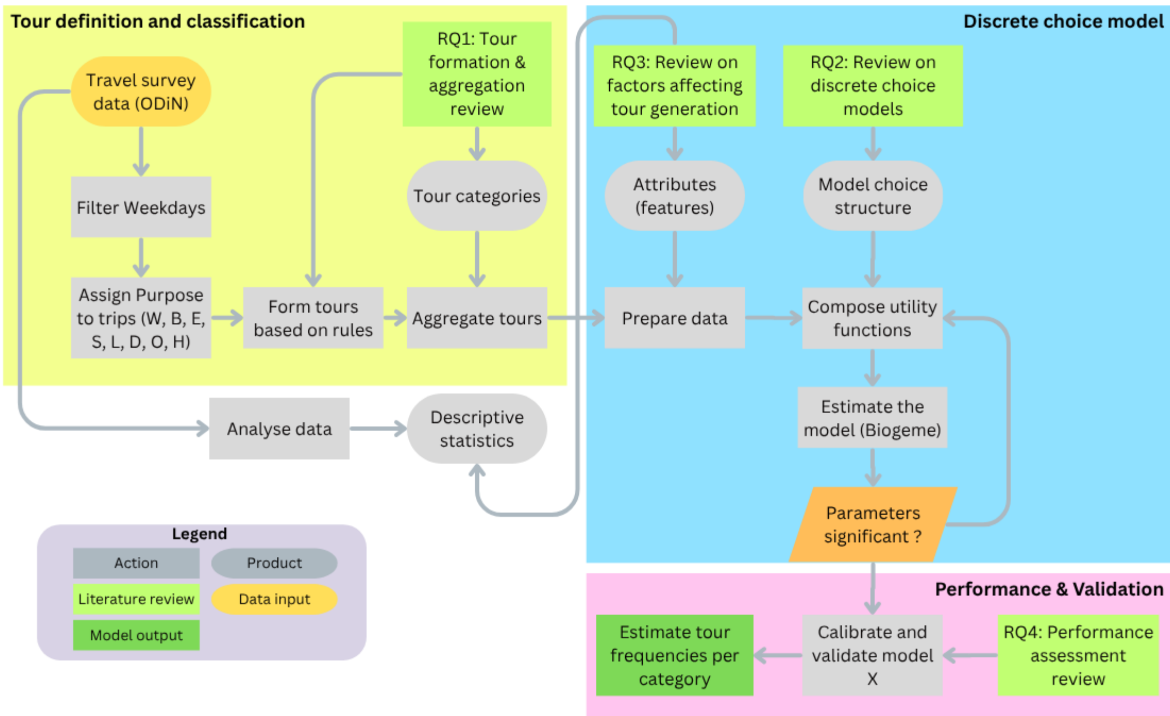


Figure 3.1: Research flowchart

3.3 Discrete Choice model

This section addresses the model development phase (blue area in Fig. 3.1). Discrete choice models will be used to estimate the frequencies of different types of tours by modeling the choice process of individuals for daily travel decisions. Many different choice structures and model types can be used for designing the model (Subsection 3.3.2). A discrete choice model requires alternatives to replicate the decision process of a traveler. As the end goal is to estimate the rates of different trip-chains, the model structure and choice set can become very complex. Therefore, to reduce the complexity for estimation purposes and keeping in mind a few principles from Bowman and Ben-Akiva (2001), the trip chains are first classified based on their primary activities (6 primary purposes) and a set of alternatives is defined for choosing different daily patterns with combinations of these categories or staying at home. However, this step does not model the choice between trip chains yet. Therefore, the choice process will be modeled in two stages, see Fig. 3.2. In the first stage, a person chooses a daily pattern between staying home, working from home or undertaking at least a tour out of the primary purpose alternatives (work, shopping, etc., or combinations), see section 3.3.1 for a detailed explanation of alternatives. In the second stage, the choice is made for the number and type of trip-chains of a certain primary purpose category previously chosen in stage 1. A separate model will be estimated for trip-chains of each primary purpose in the second stage sub-models (work, education, shopping, leisure, other) as it is expected that different tours are generated by different circumstances (see Section 2.3.2), and thus the parameters that might be significant in predicting tour rates can vary per tour type. Utility functions are determined for each alternative that the decision-maker can choose from (Subsection 3.3.4). Explanatory attributes can be socio-demographic or zonal variables that affect the choice of making different tours (see Section 4.4).

Model estimation is performed using travel diary (revealed preference data) and socio-demographic data from the National Travel Survey (ODiN). It outputs the parameters for different attributes, which can be used to predict the tour patterns for a given population input. Interpretability and transferability are strong qualities of such models as it is easy to trace where a decision comes from and interpret the influence of certain attributes.

3.3.1 Alternatives

A discrete choice model consists of alternatives that the decision-maker can choose from based on utility functions and utility maximization theory. The chosen and non-chosen alternatives in the models consist of tour patterns (how many and which types of tours a person does) in a weekday. The model has a two-stage structure with separate sub-models, see Figure 4.2 with some example alternatives. In the first stage, a choice is made whether a person stays home (base alternative with 0 utility), works from home or makes at least one tour from the available primary purposes. For example, if a person makes only work tour(s) in a day, the Work alternative is chosen, or if he makes at least a work and a leisure tour (at least one tour from each), the Work & Leisure alternative is chosen. In the second stage, the number and trip chain for each primary purpose tours that were chosen in the first stage are determined.

There will be a separate model in the second stage per primary purpose and the set of alternatives (trip chains from each primary purpose) for each of them will be determined by analyzing the available travel survey data from ODiN. The types of tours (trip chains) that can appear from the data can be very large, but a set of significant alternatives needs to be determined from the tours that appear by filtering only the tours that appear above a threshold of the observation number for each submodel. The threshold is set to retain at least 95% of the total dataset in order to focus the model on the most frequent and representative alternatives to model average travel behavior. Excluding alternatives that occur less frequently, ensures that the number of alternatives is acceptable and enough observations are available to estimate their parameters, making the model robust while still capturing the vast majority of the observed choices. Therefore, the rare tours are excluded from the model estimation as outliers and not having enough data to estimate them.

The base alternative in the primary purpose models is the simplest tour in each of them (e.g. H-W-H for the work model, H-L-H for the leisure model etc.), so the parameters are estimated for making a tour different from those simple trip chains. Alternatives in the primary work tour model (Stage 2) might contain only one work tour per day (e.g. H-W-S-H) or multiple work tours (e.g. 2 x H-W-H tours). Therefore, the choice process starts in the first stage, in which a person chooses a daily pattern, for example Work & Leisure pattern. After that the choice proceeds to the second stage Work model and second stage Leisure model, where a trip chain (or multiple) is chosen for each of those primary purposes (e.g. H-W-H for the work purpose and 2 x H-L-H for the leisure purpose).

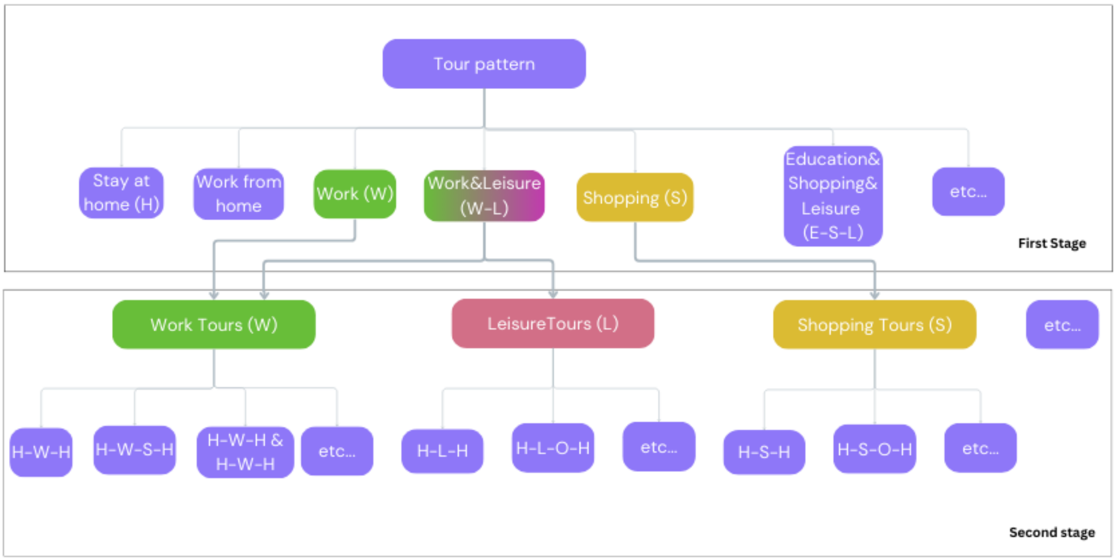


Figure 3.2: Alternative examples at each stage (Multinomial Logit Structure)

3.3.2 Logit model structures

Different logit model types will be tested and compared to determine the best model structure for estimating tour frequencies. Multinomial logit (MNL) will be the initial prototype model structure with all alternatives (see Subsection 3.3.1) in the first stage of tour pattern choice and the second stage of trip chain choice considered independent and mutually exclusive as in Fig. 3.2 . Multinomial logit models assume the independence of irrelevant alternatives (IIA property) excluding correlations between alternatives (Kenneth E. Train, 2009).

Nested logit structure, on the other hand, might be more appropriate if some alternatives are correlated and can logically be included in a nested structure. Therefore, similar alternatives can be grouped into nests and tested if the nest is significant and if it improves the overall model quality compared to the multinational logit model. For the first stage model, an alternative nested structure will be tested and evaluated with a choice between staying at home (nest 1) with two alternatives (staying at home and working from home) and going out (nest 2) with all other alternatives that involve making at least one tour, presuming that the alternatives including and not including travel are correlated within each group, see Fig. 3.3. For the second stage models (primary purpose models), the alternative choice structure will include two nests: one with one-tour alternatives and one with two or more tours (2+) as most of the alternatives contain only one tour and a few of them more tours of the same type in a single day, see Fig. 3.3. The alternative nested structures will be compared with the simple multinomial logit structures to choose the best-performing for each model.

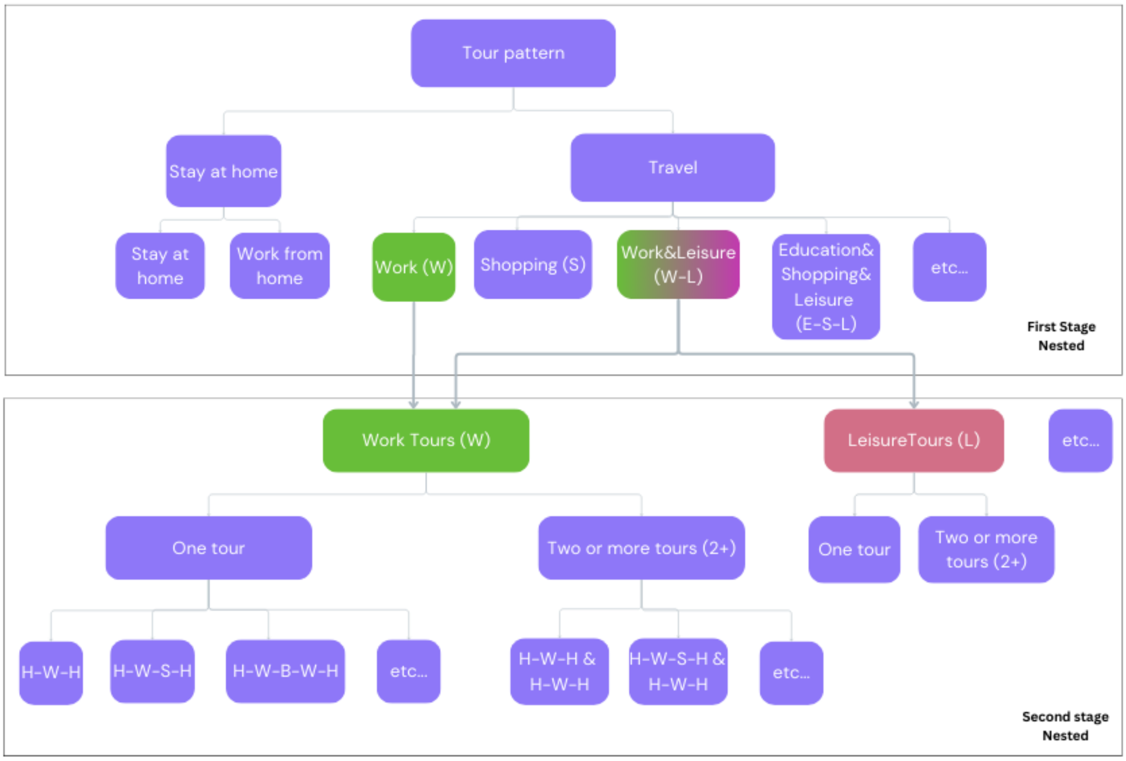


Figure 3.3: Nested choice structure with example alternatives

3.3.3 Trip chain shortening

For the second stage models, the number of trip chains (alternatives) that appear in the data is very large and some observations are filtered out as explained in Subsection 3.3.1 . To reduce the amount of data that is excluded, a trip chain shortening is carried out for the trip chains that fall outside this threshold aiming to get them within the threshold with a transformation that does not affect the core of the trip chain and visited activities. The trip chains that have consecutive repeating activities are transformed by removing one of the repeating activities (e.g. H-W-W-S-H to H-W-S-H) and the P-factor (weight of the observation) of that observation is increased with a factor as in formula 3.1 to compensate for the reduced trip chain.

$$P_{\text{factor, modified}} = P_{\text{factor}} \cdot \frac{n_{\text{initial\_activities}} - 1}{n_{\text{final\_activities}} - 1} \tag{3.1}$$

3.3.4 Utility functions

Existing literature and available data will support the choice of factors that might influence travel behavior and activity patterns. There are multiple factors affecting the activity participation of people with socio-demographics (e.g. gender, income, household, culture etc.) expected to play a crucial role in defining the preferences, constraints, and priorities that shape their travel behavior. Accessibility is also expected to have an effect on the way people organize their travel, given that the availability of modes, residential location, and activity density directly affect the feasibility and convenience of accessing different activity locations. Therefore, the selection of factors for this study will prioritize those that are commonly identified in existing research as significant predictors for travel behavior, ensuring alignment with data availability and study's objectives.

The utility functions of alternatives are composed of an alternative-specific constant (ASC) that captures the effect on the utility of an alternative that is not explained by the attributes, and personal attributes that might have an effect on the tour choice of individuals such as household size and income, employment, age, gender, car ownership or urbanization level of residence zone, etc. The available data on these attributes from the ODiN dataset will be used to test their significance in explaining tour choices in different models of daily pattern (1st stage) and primary purpose (2nd stage). The personal attributes will be included in the utility functions as dummy variables (0/1) that determine if a person makes part in that category or not, see Formula 3.2 for an example utility function of a H-W-S-H tour. The categories of personal and zonal attributes will be determined in Chapter 4.

The parameters that quantify the effect of attributes in the initial models will be the same for all alternatives, so no alternative-specific parameters. This means that all parameters are estimated to make a choice different from the base alternative (all parameters are 0). However, if differentiating effects between alternatives are not captured well during model validation, the inclusion of alternative-specific parameters in the utility functions is introduced.

$$U_{\text{H-W-S-H}} = \text{ASC}_{\text{H-W-S-H}} + \beta_1 \cdot \text{Age}_{35\text{–}64} + \beta_2 \cdot \text{Work}_{\text{FT}} + \beta_3 \cdot \text{Income}_{\text{High}} + \beta_4 \cdot \text{ZoneLevel}_1 \quad (3.2)$$

Where:

- $\text{ASC}_{\text{H-W-S-H}}$ is the alternative-specific constant for the H–W-S-H tour,
- $\text{Age}_{35\text{–}64}$ is a dummy variable equal to 1 if the individual is aged 35–64, 0 otherwise,
- Work_{FT} is a dummy variable equal to 1 for full-time workers,
- $\text{Income}_{\text{High}}$ is a dummy variable for individuals in the high-income category,
- ZoneLevel_1 is a dummy variable equal to 1 for residents in Zone Level 1 (ZL1).

3.3.5 Estimation strategy & Software

The estimation strategy for the models is backward elimination, starting with all available attributes and removing insignificant parameters one by one and re-estimating the model until all parameters are significant. The significance of parameters is tested per model using the robust t-test at a 95% confidence level (Eq. 3.3). The parameter to be removed is chosen based on correlation with other parameters. The highest correlated pair from the correlation matrix is identified, and the parameter with the highest p-value of those two is removed in the next iteration. The estimation of the models is performed using Biogeme, an open-source software package specialized for the estimation of discrete choice models that can be integrated in Python (Bierlaire, 2016). Biogeme offers capabilities for model specification, statistical outputs, and handling of large datasets, making it well-suited for this research.

$$|\text{robust } t\text{-test}| > 1.96 \Rightarrow \text{significant parameter with 95\% confidence} \quad (3.3)$$

The model estimates the parameters to maximize the log-likelihood function (Equation 3.4, where n represents each person choosing alternative j(n)), so maximizing the probabilities of choosing the observed alternatives. The ODiN travel survey aims to be representative of the whole population by surveying people from different groups. For each respondent a P-factor is given, representing the number of people in the population that each person represents (see Section 3.1). Therefore, the P-factors (w_n) are used to scale the effect of an observation in the log-likelihood function. A weighted log-likelihood function is then used as in formula 3.4, in which observations from individuals that represent a bigger group in the population have a higher weight in the function and vice versa to make the model representative of the whole population (Ortelli et al., 2023).

$$\ln L(\beta) = \sum_{n=1}^{N} w_n \ln P_{n(i)} \tag{3.4}$$

Where P(i) for Multinomial logit:

$$P(i) = \frac{e^{V_i}}{\sum_j e^{V_j}} \tag{3.5}$$

And P(i) for i in nest m in Nested logit:

$$P(i) = P(i \mid m) \cdot P(m) = \left(\frac{e^{\mu_m \cdot V_i}}{\sum_{i' \in m} e^{\mu_m \cdot V_{i'}}} \right) \cdot \left(\frac{e^{\mu \cdot IV_m}}{\sum_{m'} e^{\mu \cdot IV_{m'}}} \right) \tag{3.6}$$

where Inclusive Value of nest m is:

$$IV_m = \frac{1}{\mu_m} \ln \left(\sum_{i \in m} e^{\mu_m V_i} \right) \tag{3.7}$$

3.4 Performance & Validation

The methods for validating and evaluating the performance of the developed models are specified in this section.

The ρ^2 (Rho-squared) value is a common goodness-of-fit measure in discrete choice models. It indicates how well the model explains the observed choices compared to a baseline null model (all parameters equal to 0, which assumes all alternatives are equally likely), by comparing the likelihood of both models using formula 3.8 (Train, 2002). The ρ^2 (Rho-squared) can only be used to compare two models that are estimated on the same data with the same set of alternatives. In that case it can be said that the model with higher ρ^2 explains the choices better. However, to ensure fair comparison between models with the same alternatives but different numbers of parameters, $\bar{\rho}^2$ (Rho-bar-squared, see formula 3.9) is a better measure as it accounts for differences in the number of parameters p and sample size N.

$$\rho^2 = 1 - \frac{L(\beta)}{L_0} \tag{3.8}$$

$$\bar{\rho}^2 = 1 - \frac{\left(1-\rho^2\right)(N-1)}{N-p-1} \tag{3.9}$$

The performance of each model is also tested on a validation dataset (unseen data). The available dataset for each model is divided in a training (80%) and validation (20%) subset for this purpose. The comparison between the observed choices and the predictions of the model for the validation data that the model has not seen, will be the initial in-sample validation. For each individual, a probability is calculated for choosing each alternative based on utility functions and the estimated parameters from formula 3.5 for Multinomial Logit models (Train, 2002). Furthermore, calculating the probabilities for a nested logit model is based in formulas 3.6 and 3.7 (first calculating the probability for choosing a nest based on the inclusive value of the nest and then multiplying it by the probability of choosing an alternative within the nest).

As a measure of performance, summing over probabilities of all individuals for each alternative will provide an estimation of the amount of times that each alternative has been chosen from the model to be compared with the observed frequency of each alternative, see formula 3.10. An out-of-sample validation is also carried out by testing the model using data from a different temporal setting (different years of ODiN from the ones used for training) and use again formula 3.10 to compare observed and predicted choices on the external validation dataset.

$$\hat{C}_j = \sum_{i=1}^{N} P_{ij} \tag{3.10}$$

To evaluate the overall fit of the model to observed frequencies, the total normalized absolute error is defined as the sum of absolute differences between estimated and observed counts divided by the total observed count, see formula 3.11. This metric provides an intuitive measure of aggregate model accuracy.

$$\text{Normalized absolute error (NAE)} = \frac{\sum_{i=1}^{n} |\text{Estimated}_i - \text{Observed}_i|}{\sum_{i=1}^{n} \text{Observed}_i} \tag{3.11}$$

3.4.1 Likelihood ratio test

The likelihood ratio test is used to test hypotheses and compare models. This test is utilized in testing whether an unrestricted model (more parameters or nests) is significantly better than a restricted model. The test statistic is two times the difference between the constrained and unconstrained maximum log-likelihoods, see formula 3.12 . It is compared against the Chi-squared critical value with degrees of freedom equal to the difference in the number of parameters. If it exceeds the critical value, the null hypothesis is rejected and the unrestricted model is significantly better (Train, 2002).

$$-2\left(\text{LL}(\boldsymbol{B_H}) - \text{LL}(\boldsymbol{B})\right) \sim \chi^2_k \tag{3.12}$$

3.4.2 Validation within subgroups

In addition to comparing the overall predicted and observed tour counts, the validation is also conducted within specific population subgroups, such as students, the elderly, or other relevant demographic categories. This subgroup-level analysis allows for a more detailed evaluation of how well the model captures distinct travel behavior associated with different population segments. Based on the gathered insights about capturing these variations across segments, the need for further refinement is evaluated. The refinement could include additional complexity (e.g., alternative-specific parameters) to better capture the heterogeneity in travel behavior and improve model performance.

3.4.3 Case study application

To evaluate the practical application of the developed tour generation model, a case study is conducted in which the performance of the model is compared against the current method employed by a tour-based model (Food Valley). The models are trained on the same ODiN datasets, and the comparison is performed on a test dataset (population) that has not been used during the model estimation phase of either model to ensure an unbiased evaluation. The results provide insights into whether the developed model offers improvements over the existing method in representing real-world travel behavior.

The models are not directly comparable, as there are several differences in the components of the models, specifically the personal and zonal attributes that are used, and also the type of model output. The tour frequency model from Food Valley operates at the level of personas, defined by 5 attributes: age, employment, household income, car ownership, and zone level, see Table 3.1 for the categorization of attributes. The calculation of frequencies is simply based on the occurrences of each tour within each Persona group. The tours are aggregated for each persona and the frequency is derived from dividing the tour weight (sum of population segment factors (Factor P) of all people in the persona group that made that tour) by the persona weight (sum of factors (Factor P) of all individuals in the persona group, whether they made that tour or not), see formula 3.13.

The output of the tour frequency model is a frequency per tour type per persona, in contrast to the developed discrete choice model, which outputs probabilities for each tour. Therefore, comparing model results directly is not straightforward. Because of that, applying both methods in a test dataset that is different from the training dataset is a suitable approach for evaluating their performance. A full set of tours will be generated from both models, and they will be compared with the tour occurrences from ODiN to evaluate how well each model replicates the observed counts for different tours. However, the set of tours that is modeled with the discrete choice model is slightly different from the tour frequency model of the Food Valley project due to following different methods for eliminating rare trip chains from the set of modeled tours. Therefore, only the tours that are modeled from both methods will be compared.

$$f_{t,p} = \frac{w_{t,p}}{w_p} \cdot \alpha_p \tag{3.13}$$

Where:

- $f_{t,p}$ is the frequency of tour type t for persona p
- $w_{t,p}$ is the total weight of people in persona p who made tour t
- w_p is the total weight of all people in persona p
- α_p is the correction factor for shortened tours (persona increase factor)

Table 3.1: Explanatory attributes used in the Food Valley model

Attribute	**Categories**				
Age	0-17	18-34	35-64	65+	
Employment	Unemployed	Part-time	Full-Time		
Household income	0-30 k	30k-50k	50k+		
Car Ownership	No Car	1 Car	2+ Cars		
Zone Level	ZL1	ZL2	ZL3	ZL4	ZL5

4 Data Analysis and Preparation

In this chapter, a thorough analysis of the available ODiN data from 2018 to 2023 is conducted. The purpose of this analysis is also to provide insights into the factors that can influence tour generation and support modeling choices, which will form the foundation for the construction of the tour frequency model. The formation of tours from trip data and variations in travel trends throughout the different years of data are discussed in section 4.1, supporting the suitability of the datasets for the modeling process. Some general trends of the data that is used for modeling are discussed in section 4.2. The sets of alternatives for the models that are developed are defined in section 4.3. By applying descriptive statistics to the tour dataset in section 4.4, the aim is to better understand key trends and patterns in the daily travel behavior of different population segments. This understanding is crucial for identifying potential relationships that may affect the frequency of different types of tours within various population segments. The general format of the input data is described in section 4.5.

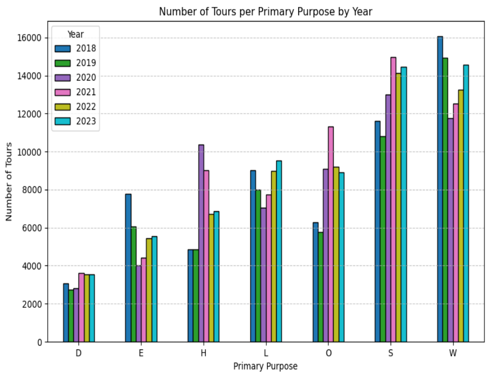
4.1 Tour Formation and trends

The combined dataset of 6 years (2018-2023) from ODiN contains reports of the daily movements of each person who responds to the survey. Socio-demographic and zonal information is provided for the respondents apart from travel records to link their travel behavior to their background. The trip records are processed to form tour sequences that will be used for data analysis and model construction. Tours are generated as sequences of home-activities-home (joining activity purposes) for the reports of each individual ID present in the data. Therefore, returning at home is the indication for ending a tour and starting a new one if more trips are available for the same person. The tours are forced to start and end at home, so if only outgoing trips (not coming back home) are registered for a person, the home destination is added at the end of the sequence to complete the tour (e.g. a home-work-shop is transformed into home-work-shop-home tour). The primary purpose and the secondary activities are defined for each tour based on the hierarchy defined in the section 3.2. The processed dataset resulted in a total of 315530 tours for 248181 respondents, while 42649 additional respondents reported that they stayed at home for different reasons (no trip records).

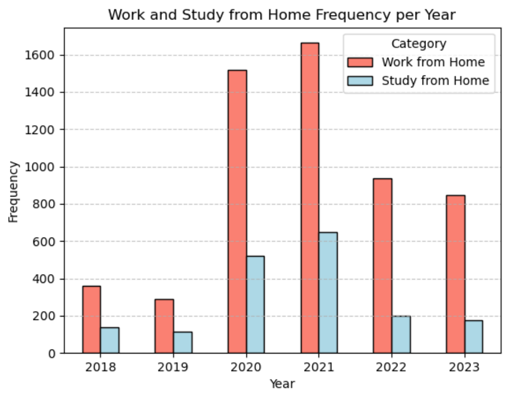
An analysis of yearly tour data from 2018 to 2023 reveals significant shifts in travel behavior, particularly during and after the COVID-19 pandemic, see Fig. 4.1a. The largest shifts occurred in 2020 and 2021, when a significant increase was observed in the number of individuals staying at home, as well as in shopping and other tour categories. On the other hand, work, education, and leisure-related tours experienced a substantial decline during the same period. These trends align with the lockdowns due to the pandemic and remote work & study policies, see Fig. 4.1b for the study and work from home occurrences through the years.

After the pandemic, from 2022 onward, travel behavior has shown signs of returning to pre-pandemic levels. However, work and education tours have not yet fully recovered to their 2019 levels, likely due to the continued policies of hybrid work and remote learning options. On the other hand, shopping and other discretionary tours have remained high compared to pre-pandemic years, suggesting a potential shift in travel behavior or lifestyle changes. The percentage of individuals staying at home has been decreasing since 2021, indicating a recovery towards normal mobility activity.

Given the significant anomalies in travel behavior during 2020 and 2021, these years are excluded from model development to ensure that the estimated relationships between explanatory variables and tour choices reflect long-term behavioral patterns rather than temporary disruptions. Including these years could introduce biases, as the data during this period do not represent typical travel demand. Furthermore, even though the behavior has more or less normalized, the effect of the pandemic might have shifted some aspects of travel behavior permanently as more people have normalized working from home and shopping activities have remained high. Therefore, focusing on the recent 2022-2023 data for training and validation will lead to a more reliable and representative model for predicting tour frequencies. The data from 2018-2019 can be used as an external validation dataset for testing the model in a different temporal setting.



(a) Variation of tour occurrences through years



(b) Variation of Work and Study from Home categories through years

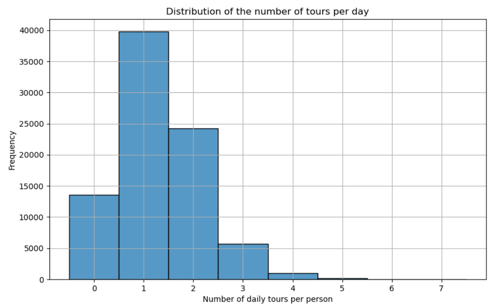
Figure 4.1: Trends in the available data between 2018-2023

4.2 Tour Patterns

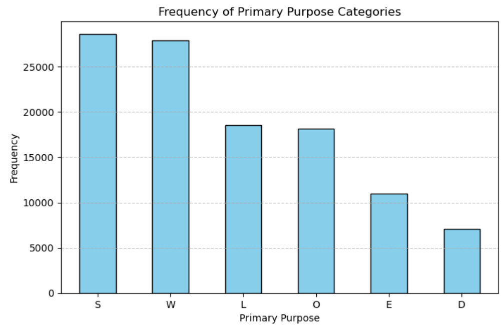
The tours from 2022-2023 that will be used for model training and validation are analyzed to identify some trends to support modeling choices. The dataset consists of 111089 tours, and the distribution of the number of tours that travelers reported in a day is shown in Fig. 4.2a with an average of 1.13 tours per day. One-tour daily patterns are the most occurring in the dataset with around 47% of the respondents being part of that group, followed by two-tour patten with 29% and no-tour (staying at home) pattern with 16%. Patterns with up to three daily tours cover 99% of the data, indicating that undertaking more than three tours within a day is very rare. Figure 4.2b shows the distribution of tours across different primary purposes. Primary work and shopping tours have the highest share, followed by leisure, other, education, and escort tours, decreasing in share in that order.

The frequencies of different numbers of intermediate stops (including primary purpose) for the tours and a cumulative distribution are shown in Fig. 4.2c. Around 76% of the tours only visit the primary destination (1 intermediate stop) and return home, while 17% and 5% visit one and two secondary stops respectively. These three categories make up for around 98% of the observations, indicating that tours with more than three intermediate stops (2 secondary stops) are very rarely observed.

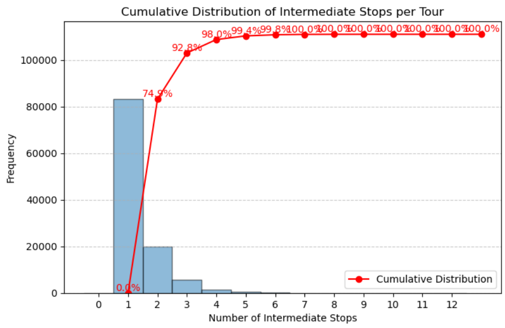
The percentage frequency plot in Fig. 4.2d shows the share of each primary purpose for each number of intermediate stops to identify the types of tours that are more likely to visit only the main destination or be combined with secondary activities. As can be seen, escort, education, leisure, and other primary purpose tours visit usually the primary activity or combine it with one additional stop, with very rare cases visiting more than one secondary stop (2+ intermediate stops). On the other hand, work and shop tours are more likely than the other primary purposes to include more than one secondary stops, but still the majority of tours fall under the one and two-intermediate-stop categories.



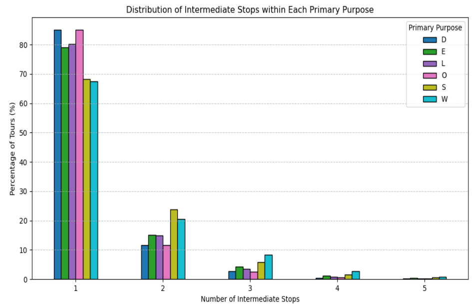
(a) Frequencies of daily number of tours



(b) Frequencies of primary purpose categories



(c) Distribution of intermediate stops



(d) Distribution of intermediate stops within each primary purpose

Figure 4.2: Tour data patterns

4.3 Filtered alternatives

The number of different trip chains that appear in the dataset of formed tours is very large, and not every trip chain can be modeled as the choice set becomes enormous (see Section 3.3.1). Therefore, a significant choice set has been defined for each primary purpose model based on the occurrence of trip chains in the data to retain at least 95% of the data (97% retained) while keeping the number of alternatives acceptable for capturing the majority of the tours. The number of filtered alternatives and respective data counts for each primary purpose model are shown in Table 4.1. The trip chain patterns (can also be two or more tours in a day) that make up the alternative sets for each model in the second stage are shown in Table A.1 of Appendix A. The combined number of alternatives of all second-stage models adds up to 158, with 119 trip chains (the remaining 39 alternatives are combinations with two or more of these 119 chains).

Table 4.1: Alternative counts for second stage models

Primary purpose	**Alternatives**	**Data count**
Work	47	25302
Education	22	10159
Shop	42	23600
Leisure	22	16892
Escort	18	4975
Other	7	16982

4.4 Explanatory attributes

Possible explanatory attributes that can have a significant effect on tour generation, which can be retrieved from ODiN as personal or zonal attributes, but also supported by literature in Section 2.4 are analyzed in the plots in Fig. 4.3 which show the variation in tour rates of different primary purposes for different categories of the given attributes.

Household income is a common predictor for tour generation as it is expected to affect the type of activities that individuals from households with different incomes do. The income of households in ODiN is represented by percentage groups (0-100). Given that information, three income groups are defined: Low (0-40% income groups), Middle (40-80% groups) and High (80-100% groups). The tour rates of primary purpose categories by income group are shown in Fig. 4.3a. The high variation (higher slope of the lines) in some categories indicates a possible significant effect of the predictor on tours of that primary purpose type. Work (W) tours seem to have the biggest variation in rate from different income groups, indicating that higher-income households undertake more work tours. The stay-at-home (H) alternative and shopping tours (S) seem to increase sharply for households with low income.

The occupation of individuals can also be crucial in predicting tours of certain types. The occupation is categorized as people that work part-time (PT), full-time (FT), don't work or study (NW) and students (St). The tour rates per category are shown in Fig. 4.3b, where work (W), shopping (S), and education (E) tours seem to have a lot of variation by the occupation of individuals.

Age is also explored as a possible predictor with four age categories (0-17, 18-34, 35-64, 65+). As can be seen from Fig. 4.3c, education (E) and leisure (L) rates seem to decrease when getting older, while shopping (S) tours increase. The work tour rate increases up to the third category but sharply decreases for people in the 65+ category mainly due to retirement.

Gender can also have an impact in certain tour types and is supported by literature as an important factor in tour generation. The tour rates by gender are shown in Fig. 4.3d, where can be seen that men take more work tours and fewer shopping tours compared to women. Females also have a higher rate of drop-off (D) tours, indicating that the gender might have an influence in generating escort tours.

The size of the household might be a significant attribute in estimating certain tours. Three size groups are considered: Single (1 person), Medium (2 people), and Large (3+ people). There appears to be a sharp increase in education (E) and escort (D) tour rates in the large households, most likely due to the presence of children. On the other hand, the shopping tour rate seems to decrease with increasing household size, probably due to sharing duties and household members shopping for the whole family.

Car ownership is also a common predictor for generating tours as it affects the mobility access of a household. The car ownership attribute is categorized in: no car (NCA), one car (CA1) and at least two cars (CA2). Car ownership seems to increase the rate of work, education, and leisure tours, but decrease the rate of shopping tours (Fig. 4.3f).

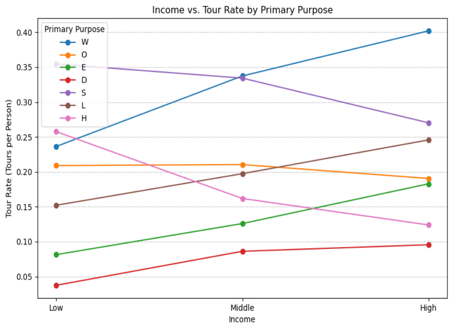
The urban level of the home address is also provided for each individual in ODiN with a rating from 1 (Strongly Urban) to 5 (Not Urban). It considers the number of addresses per postcode to classify the level. The variation of tour rates within these urban levels are shown in Fig. 4.3g. As can be seen, people living in rural areas undertake more work tours and fewer shopping tours compared to those living in urban areas. The reason could be that rural residents combine work with more activities (like shopping) and take longer tours as they live far from the activities, resulting in less primary shopping tours and more primary work ones.

The zone level is another zonal attribute calculated from a combination of the number of households and public transport stops in a postal code zone (PC4). The households and public transport stops are considered in a radius of 2 km around the PC4 area. A factor (x125) is used to scale the public transport stops to the magnitude of households ($125 \times \text{OV stops/km}^2 + \text{HH/km}^2$ with a radius of influence of 2 km). This attribute aims to quantify the level of activity and public transit accessibility in the residential location of a person. It also has a scale from 1 (High level) to 5 (Low level). The tour rates of different primary purposes per zone level seem very similar to the Urban level attribute, which suggests a high correlation between the two.

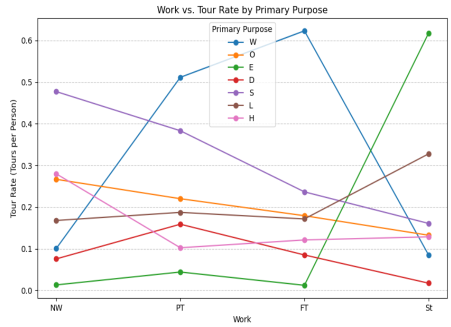
The final set of attributes and categories for each of them, that will be tested in the developed models is shown in Table 4.2.

Table 4.2: Set of attributes and categories tested in the models

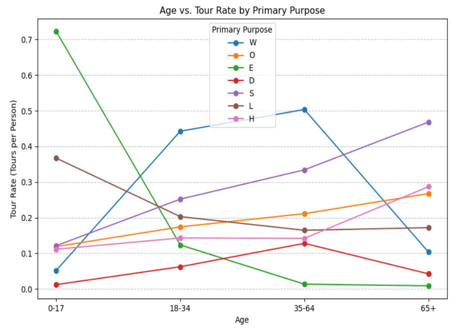
Attribute	**Classes**				
Age	0-17	18-34	35-64	65+	
Gender	Male	Female			
HH income	Low (0-40 %)	Medium (40-80%)	High (80+ %)		
Occupation	Unemployed	Student	Part-time	Full-time	
HH size	Small (1)	Medium (2)	Large (3+)		
Car ownership	0	1	2+		
Urban level	1	2	3	4	5
Zone level	1	2	3	4	5



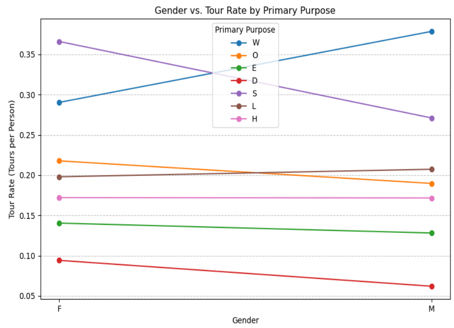
(a) Household income



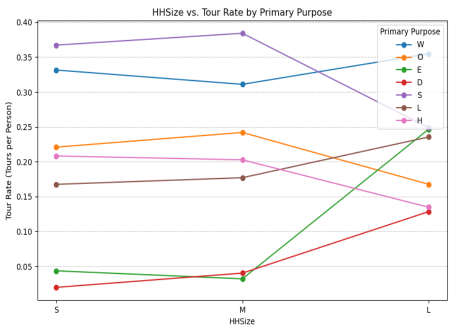
(b) Occupation



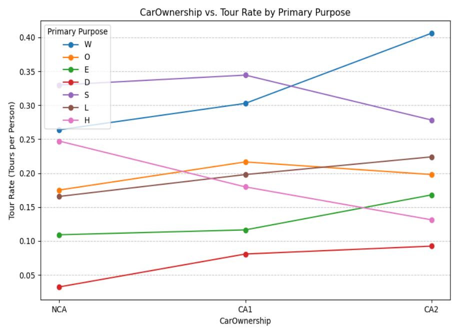
(c) Age



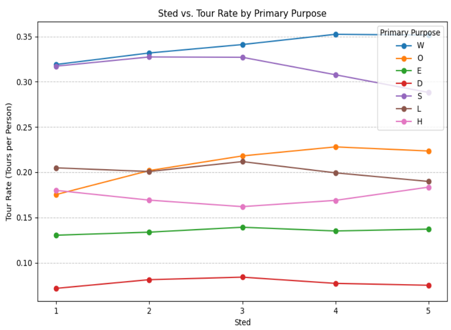
(d) Gender



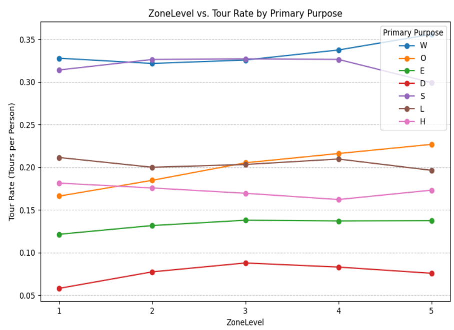
(e) Household size



(f) Car Ownership



(g) Urban level



(h) Zone level

Figure 4.3: Tour rate variation per attribute

4.5 Data preparation

The estimation of a discrete choice model requires transforming the data in the required format. For each observation (person) in the filtered dataset, a set of attributes are assigned that determine which category of the attributes that were suggested in section 4.4 the person is part of. For each part of the model, the choice of that individual for one of the alternatives is also given. Each row of the training data consists of the person ID, factor P, which is the weight of that person in the log-likelihood function (see section 3.3.5), attributes as binary variables per category, and the choice of alternative (Pattern). For the model in the first stage, the alternatives consist of daily patterns of what primary purpose tours an individual chooses to do in a day, while the alternatives of the models in the second stage consist of trip chains that people can choose for those tours for each primary purpose. For an example of an input row for the work model that chooses the trip chain see Fig. 4.4.

imiter:

	OPID	FactorP	Year	TourCount	Tour	Tour_2	Tour_3	Age_1	Age_2
1	55834100372	229	2018	1	H-W-W-B-H			0	1

elimiter:

	_1	Age_2	Age_3	Gender_1	HHSize_1	HHSize_2	Income_1	Income_2	Work_1	Work
1	0	1	0	0	0	0	0	1	0	

ter:

	Work_2	Work_3	CarOwnership_1	CarOwnership_2	Sted_2	Sted_3	Sted_4	Sted_5	Pattern
1	1	0	1	0	0	0	1	0	H-W-W-B-H

Figure 4.4: Estimation example input

4.6 Conclusion

This chapter provided the foundation for the model development by exploring the patterns of tours, analyzing trends in travel behavior across 2018–2023, and motivating the selection of 2022–2023 data for estimation. Descriptive statistics on the selected data revealed key insights into tour patterns and their variation across population groups. The selection of alternatives for each primary purpose is crucial for model specification. Finally, a set of explanatory attributes with theoretical and empirical relevance is chosen to be included in the model estimation process to explain choices made by individuals. This data analysis supports the specification of the models that were defined in Chapter 3 for estimation, leading to the results that are presented in Chapter 5.

5 Results

This chapter presents the outcomes of the estimated models and evaluates their performance to provide an answer to the main research question about which model structure and predictors should be used for tour generation. The results are presented for each part of the model (the first stage and primary purpose submodels) as defined in the Methodology chapter. For each submodel, the estimated parameters are examined to identify significant predictors that influence tour choices, so addressing the research question regarding relevant explanatory variables. Model performance is assessed using the evaluation metrics defined in Section 3.4, allowing for a comparison between the developed model structures. These comparisons support the main research question on providing the highest explanatory power for predicting tour frequencies.

The first step of the analysis involves analyzing the estimated parameters and performance metrics (log-likelihood and $\bar{\rho}^2$) from Biogeme output. Secondly, the base Multinomial Logit (MNL) structure of each submodel is compared with an alternative Nested Logit (NL) structure as defined in Section 3.3.2. A likelihood ratio test is performed to determine whether the nested model provides a statistically significant improvement, guiding the selection of the most suitable model structure for tour prediction. Thereafter, the validation of the models is carried out by comparing the predicted and the observed choices in a validation dataset made of 20% of the ODiN data from 2022–2023, which was excluded from the model training phase. The accuracy of the model predictions is assessed using the normalized absolute error (NAE) to determine how well the estimated models reflect real-world observed choices as explained in section 3.4. Additionally, the generalizability of the models is tested on an external dataset from 2018–2019 to examine their performance in a different temporal context. The final validation step is the assessment of the model performance within subgroups (e.g. students) to detect any possible anomalies and suggest possible improvements in the model specification. Finally, the validated model is compared with a current method used for generating tours in a case study to evaluate their performance on unseen data.

The results provide insights into the effectiveness of the discrete choice model in capturing daily activity patterns and evaluate how including various personal and zonal attributes improves predictive capability.

5.1 First stage daily pattern results

The first stage model has been estimated with all attributes as defined in section 4.4 with a set of 34 alternatives: stay at home, work from home and 32 combinations of different types of tours that a person can do in a day (e.g. S means at least a shop tour, W-D-O means at least a work, drop-off and other tour). The parameter estimation results for the multinomial logit and nested logit (2 nests, stay at home and travel) model structures are shown in Table 5.1. The base alternative is staying at home (alternative 9 with utility 0) and all the parameters are estimated for choosing a pattern different from that. Only the significant parameters are shown, as the non-significant ones were removed during model estimation.

The set of estimated parameters contains the beta parameters for each attribute class and an ASC for each alternative, which is a constant that indicates the preference for a certain alternative relative to staying at home (base alternative), regardless of the characteristics of the person making the choice. The ASC value quantifies the behavioral preference that is not explained by the explanatory attributes. It can be seen that patterns that occur quite often in the data (especially the single tour type patterns, e.g. W, S, O, etc.) have a higher ASC (less negative) compared to alternatives that occur less often. This means that in a set of individuals with the same personal characteristics (age, gender, occupation, etc.), the patterns with higher ASCs have a higher probability of getting chosen. As expected, tour patterns involving only work tours (W) have the least negative constant (-0.57 in MNL, -0.47 in NL), followed by patterns involving only shopping tours (S), indicating a relatively high preference for work-based and shopping travel. In contrast, more complex patterns involving combinations of escorting, leisure, shopping, or multiple purposes (e.g., E-D, W-D-O, S-L-O) have strongly negative constants (up to -5.7), indicating these combinations are chosen less frequently if the effect of individual-specific attributes is absent.

The results also reveal significant parameters (different from 0) for personal and zonal attributes. They show that individuals aged 18-34 and 65+ have lower utility for engaging in a tour compared to children (0-17, base), suggesting lower mobility needs, especially for seniors. Car ownership (either 1 or 2+) increase the probability of traveling compared to people that do not own a car, likely due to increased accessibility. Similarly, medium and high income groups show an increased likelihood of going out compared to low income groups, which can be explained by higher activity participation and financial capacity. People in larger households (medium or large, vs. small) are less likely to make tours, possibly reflecting coordination or household responsibility sharing. Furthermore, the occupation plays a major role: individuals with part-time ($B = 1.09$), full-time (1.02) work, or who are students (0.73) are significantly more likely to participate in activities compared to those who are unemployed. This supports the assumption that mandatory activities are a strong driver of tour formation (primary purpose assumption). Regarding urbanization, the coefficients show a slightly positive effect for traveling while living in less urban areas (levels 2–4, with base level 1 being the most urban), possibly indicating more reliance on motorized transport or less access to amenities.

The $\bar{\rho}^2$ shows improvement when introducing nests to the model (from 0.237 for MNL to 0.238 for NL). The nest parameter for the home alternatives (stay at home and work from home) is significant with a value of 3.3, indicating unobserved correlation among these alternatives. Furthermore, the log-likelihood value shows improvement in the nested logit structure, so applying the likelihood ratio test with 1 degree of freedom (one additional parameter), it can be concluded that the independence of irrelevant alternatives (IIA) property is violated, and the nested model significantly improves the model fit.

Table 5.1: Estimation results for the first stage model (Multinomial and Nested)

Parameter	Alternative	MNL	NL
ASC_0	D	-3.06	-2.96
ASC_1	D-O	-4.45	-4.35
ASC_10	L	-1.76	-1.65
ASC_11	L-D	-4.34	-4.24
ASC_12	L-O	-3.49	-3.39
ASC_13	O	-1.57	-1.47
ASC_14	S	-0.97	-0.87
ASC_15	S-D	-3.25	-3.15
ASC_16	S-D-O	-4.85	-4.75
ASC_17	S-L	-2.37	-2.26
ASC_18	S-L-D	-4.56	-4.46
ASC_19	S-L-O	-4.07	-3.97
ASC_2	E	-1.65	-1.55
ASC_20	S-O	-1.96	-1.86
ASC_21	W	-0.57	-0.47
ASC_22	W-D	-3.69	-3.59
ASC_23	W-D-O	-5.48	-5.37
ASC_24	W-E	-4.28	-4.18
ASC_25	W-L	-2.21	-2.10
ASC_26	W-L-D	-5.38	-5.28
ASC_27	W-L-O	-4.84	-4.74
ASC_28	W-O	-2.12	-2.01
ASC_29	W-S	-2.45	-2.34
ASC_3	E-D	-5.70	-5.59
ASC_30	W-S-D	-5.15	-5.05
ASC_31	W-S-L	-4.50	-4.40
ASC_32	W-S-O	-4.10	-3.99
ASC_33	Work from Home	-2.80	-1.37
ASC_4	E-L	-2.28	-2.18
ASC_5	E-L-O	-5.41	-5.31
ASC_6	E-O	-3.60	-3.49
ASC_7	E-S	-3.63	-3.53
ASC_8	E-S-L	-4.98	-4.87
B_Age_18_34	All	-0.14	-0.09
B_Age_65	All	-0.33	-0.26
B_CarOwnership_1	All	0.35	0.21
B_CarOwnership_2	All	0.40	0.21
B_HHSize_L	All	-0.30	-0.17
B_HHSize_M	All	-0.18	-0.08
B_Income_H	All	0.37	0.28
B_Income_M	All	0.37	0.26
B_UrbanLevel_2	All	0.09	0.06
B_UrbanLevel_3	All	0.10	0.03
B_UrbanLevel_4	All	0.08	0.05
B_Work_FT	All	1.02	0.94
B_Work_PT	All	1.09	0.84
B_Student	All	0.73	0.52
MU_no_tour	Home alternatives		3.31
$\bar{\rho}^2$		0.237	0.238
Log-likelihood		-181357	-181035
Norm. abs. error		0.03	0.033

The developed models are tested on the validation dataset, and the results of the observed and estimated counts for each alternative are plotted in Fig. 5.1. Both model structures seem to predict similarly and very well the observed choices in the validation data. A visible difference between the results from the two structures is the allocation of home patterns; while the total number of home patterns is almost the same from the two models, the allocation between stay at home (H) and work from home (WFH) slightly differs as the nested structure estimates less stay at home and more work from home patterns compared to the multinomial structure, likely due to accounting for correlation between those two alternatives. As can be seen from Table 5.1, the normalized absolute error of the multinomial model in the validation dataset is slightly smaller (3%) than the error of the nested model (3.3%). An additional validation is performed in an external dataset (retrospectively in 2018-2019), and the predicted choices are plotted against the observed ones in Fig. 5.2. As can be seen, the stay-at-home and work-from-home alternatives are overestimated, while work and education are underestimated, most likely due to a shift in travel behavior. As discussed in section 4.1, there is a downward trend in work and education tours after the pandemic, as more people have normalized working and studying from home, so this overestimation of the stay-at-home and work-from-home alternatives is expected. Based on the likelihood ratio test, the multinomial logit is rejected, and the nested logit structure with parameters in Table 5.1 is the best model out of the two, because the difference in normalized absolute error is deemed negligible.

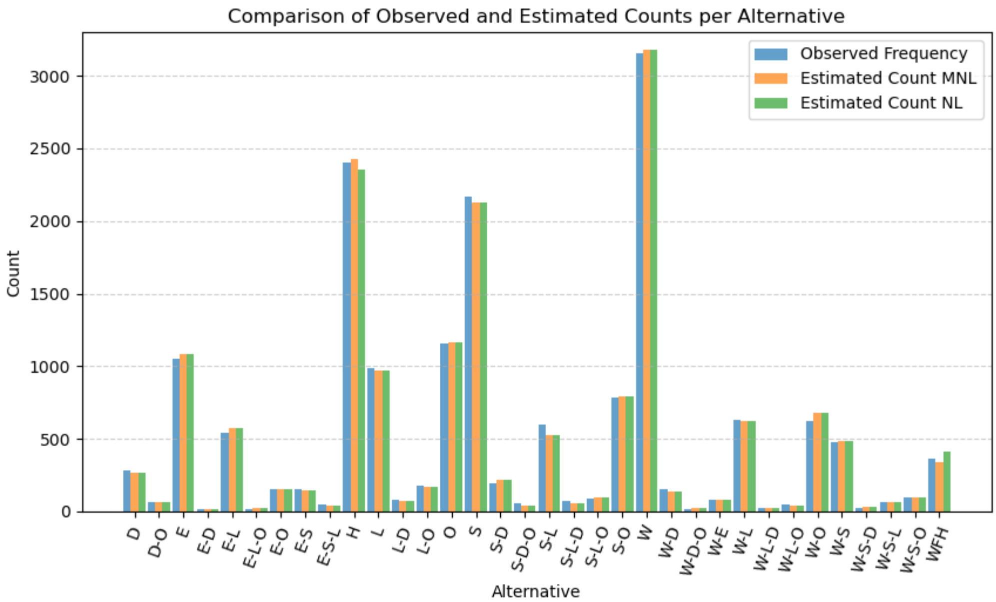


Figure 5.1: Observed and estimated counts of both model structures in the validation dataset

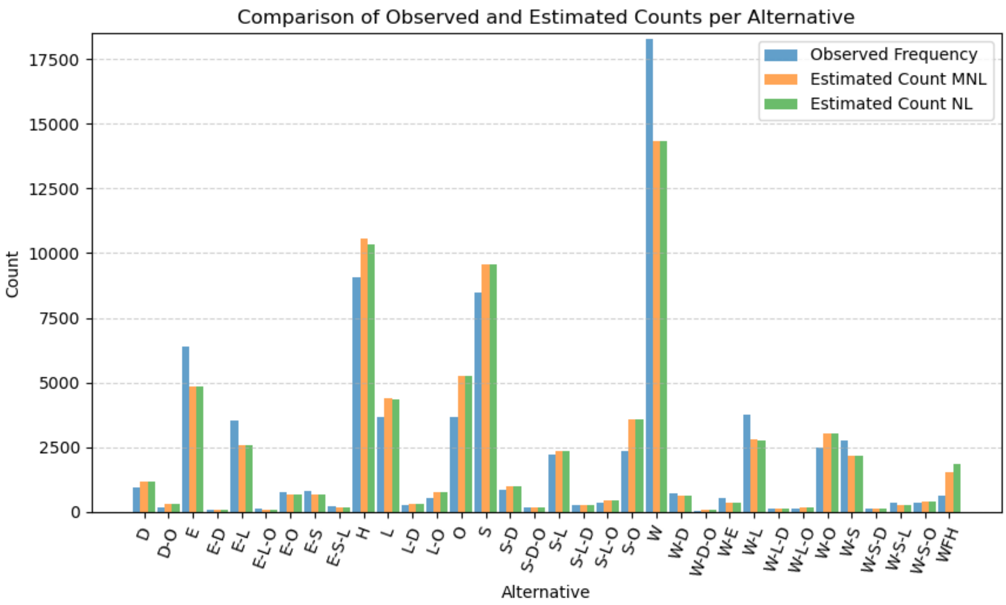


Figure 5.2: Observed and estimated counts of both model structures in the external dataset from 2018-2019

5.2 Second stage primary purpose results

The second stage models are estimated for each of the 6 primary purposes with the set of alternatives for each model as defined in Appendix A and attributes in section 4.4. The base alternative for each model is the simplest trip chain in each of them (e.g. H-W-H, H-S-H, etc.) and all parameters are estimated for choosing a trip chain different from the simple one or multiple trip chains from the same primary purpose (e.g. 2 x H-S-H or H-S-H & H-S-O-H).

For each primary purpose, the multinomial and nested structure (nest for alternatives with 2+ trip-chains) are estimated, and final estimation indicators are compared in Table 5.2. The likelihood ratio test has been applied to each primary purpose model to make the final choice for model structure, and is shown in the same table. The final significant parameters for personal and zonal attributes of each model (also the nest parameter for 2+ tour alternatives in each model) are shown in Fig. 5.3. For the full set of parameters of each primary purpose model, including ASCs, see Appendix B.

The positive parameters (blue) indicate that the utility for making a tour different than the simple one (e.g. H-W-H) is higher for that attribute. The negative parameters, on the other hand, decrease the utility of alternatives other than the shortest tour. As can be seen in general for all primary purpose models, the individuals aged 35-64, derive a higher utility for making a complex tour combined with other lower-level activities or multiple tours of the same primary purpose in a day compared to other age groups. This behavior could be due to the active lifestyle of this age group and time constraints, making them chain multiple activities in a tour. Females also have a higher likelihood in taking long tours in a few primary purpose categories (shop, education, escort), possibly due to their traditional role in the household to carry out certain tasks such as shopping and escorting kids to their activities. Car ownership increases the utility of combining shopping or other primary purpose tours with other activities, as it provides convenience (storage for shopping) and flexibility, making it easy to visit multiple locations before returning home. Household size parameters show that as size increases, the likelihood of making a complex tour in general decreases as a result of possible coordination and task distribution among household members. High income has a slightly positive impact on the utilities of the complex tours of work and education primary purposes, likely due to higher financial capacity to carry out secondary activities throughout the day. Individuals with an occupation (either working or a student) derive lower utility for making complex tours in general compared to the unemployed population. This effect might be caused by stricter time constraints for these occupied individuals. The urban level parameters have a key role in the choice of the trip chain for primary escort tours. It can be seen that living in less urban environments (other than base urban level 1) increases the likelihood of combining escort tours with other activities, possibly due to lower accessibility of these areas, pushing people to carry out other errands when escorting someone to a location for time and cost reasons. The zone level parameters were not significant for any of the models, due to correlation with urban level, and it is therefore removed as an explanatory attribute.

The estimation results in Table 5.2 show that $\bar{\rho^2}$ of the primary purpose models varies between 0.5 and 0.65, improving explanatory power over a null model. The nest parameters were significantly different from 1 for all second-stage models, except for the shop primary purpose. However, the log-likelihood did not improve for any of them over the MNL structure. Therefore, the unrestricted nested structure for all primary purpose models was rejected using the likelihood ratio test, making MNL the final structure for these models. The normalized absolute errors between model structures are also almost the same, ranging between 1.2% for other and 7.8% for escort primary purpose tours. The models fit the observed data very well for the 20% of the internal validation dataset, but there are some deviations in predicting observed choices in the external validation dataset (2018-2019), most probably due to the change in travel behavior after the pandemic. For a detailed analysis of parameters and validation plots for each primary purpose, see Appendix C.

Table 5.2: Estimation results for the second stage models (MNL & NL) and likelihood ratio test

	Multinomial logit (MNL)			**Nested logit (NL)**			**Likelihood ratio test**
Primary purpose	$\bar{\rho^2}$	**Log-Likelihood**	**Norm. abs. error**	$\bar{\rho^2}$	**Log-Likelihood**	**Norm. abs. error**	**Final structure**
Work	0.53	-36553	0.057	0.53	-36553	0.057	MNL
Education	0.65	-8713	0.057	0.65	-8713	0.058	MNL
Shop	0.52	-33664	0.051	-	-	-	MNL
Leisure	0.63	-15262	0.048	0.63	-15262	0.048	MNL
Escort	0.5	-5700	0.078	0.5	-5700	0.078	MNL
Other	0.64	-9467	0.012	0.64	-9467	0.012	MNL

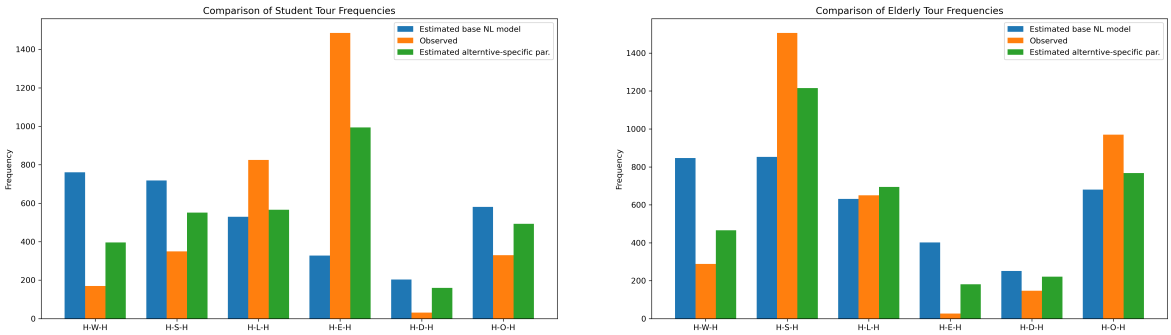
Parameters	Work	Shop	Education	Leisure	Escort	Other
Age 18-34		0.13				0.19
Age 35-64	0.20	0.21	0.31	0.12	0.16	0.48
Age 65+						
Gender F		0.15	0.21		0.30	
Car Ownership 1		0.26		-0.11		0.40
Car Ownership 2+		0.23				0.46
HH Size M	-0.10		-0.43			
HH Size L			-0.47			-0.36
Income M						
Income H	0.08		0.17			
Work Full Time	-0.14	-0.21		-0.41		-0.41
Work Part time		-0.15		-0.37		-0.22
Student	-0.48	-0.31		-0.34		-0.52
Urban Level 2					0.45	0.12
Urban Level 3					0.33	
Urban Level 4					0.49	
Urban Level 5					0.45	
Zone Level 2						
Zone Level 3						
Zone Level 4						
Zone Level 5						

Figure 5.3: Estimated parameters for the primary purpose models

5.3 Combined model validation and improvement

The combined model with two stages is tested in the 20% of unseen data to evaluate the complete model fit by comparing the estimated and observed counts of all modeled trip chains. The final normalized absolute error on the validation dataset is around 5%, and the estimated and observed counts of 20 highest observed trip chains are shown in Fig. 5.6. As can be seen, the model (blue bars) fits the observed data (orange bars) quite well. However, as stated in section 3.4, the performance of the model in specific subgroups is also evaluated to detect possible irregular behavior. The estimated counts (blue bars) of 6 simple trip-chains and the observed (orange bars) are shown in Fig. 5.4a for students and in Fig. 5.4b for elderly subgroups. It can be seen that there are significant mismatches between the counts for different trip chains. As students carry out more education and leisure tours, those are underestimated by the model, and the other tours significantly overestimated. Similarly, for the elderly, the shopping and other (O) tours are underestimated, while the other ones overestimated (e.g. work, education). Even though the model fit at an aggregate level seems very good, there seems to be compensation between underestimated and overestimated trip chains across subgroups to achieve that final fit. This means that the current model specification would not perform well in an environment where some groups dominate (e.g., a student city like Delft, where the population is not balanced as in the national level).

The good performance of the model at an aggregated level and poor performance in specific subgroups suggests that the model could be underspecified, highlighting the dominance of alternative-specific constants (ASCs) in determining the overall fit. The parameters for the personal attributes in the first stage model are the same for all alternatives and are estimated to choose a pattern different from staying at home. However, no distinction is made in the effect of the elderly (65+) between choosing a work or a shop pattern for example. The first-stage model is responsible for allocating the tours to the primary purposes and can be adjusted to account for variation between the effects of the parameters across different alternatives.



(a) Students

(b) Elderly

Figure 5.4: Comparison of simple tour frequencies within subgroups

5.3.1 First stage model respecification

For modeling the differences in the effects of attributes in different patterns, alternative-specific parameters are defined for the 6 single tour type patterns (W, E, S, L, D, O) and work-from-home (WFH) in the first stage as those are the most popular patterns and contain a single category of primary purpose allowing the parameters to distinguish the effect on each primary purpsose. For the remaining pattern alternatives (combinations of these primary purposes, e.g. W-L), the parameters are still the same for all of them to avoid overspecifying the model.

The estimated parameters for the attributes of the first-stage model with multinomial logit structure and alternative-specific parameters for the 6 single-tour type patterns and working-from-home are shown in Fig. 5.5 (see Table B.7 in Appendix B for the full set of parameters including ASCs). It can be seen that being an adult increases the likelihood of doing only work, only shopping, or only escort tours compared to being a child. Young (18-34) and middle-aged adults (35-64) are significantly more likely to work from home. As individuals get older, the utility of choosing an education only pattern sharply decreases, while the utility of leisure or complex patterns (multiple tour types) slightly decreases as people tend to have less time for extra activities, especially in a normal working day, for which this model is designed. Females derive less utility from work-only (also from home) or leisure-only patterns, but higher utility than males for shopping-only patterns, likely due to a higher preference and responsibility-taking for shopping activities. Car ownership increases the utility of all patterns that involve going out of home as it increases the accessibility and lowers travel resistance. Increasing household size decreases the likelihood of choosing an out-of-home alternative in general, possibly due to responsibility sharing, except for escort patterns for large households (3+ members), which can be due to the presence of children and dropping off/picking them up from activities. Medium and high income groups are in general more likely to choose a pattern involving traveling, as a result of additional occupations or more financial capacity to carry out activities. Working full-time or part-time as expected significantly increases the likelihood of choosing a work-only and or work-from-home pattern, and being a student or part-time worker, the likelihood of an education pattern. Individuals with an occupation derive higher utility for leisure-only or complex patterns, and lower utility for escort-only patterns, compared to unemployed groups. Living in less urban areas, the utility of choosing a work-only pattern increases, possibly due to the tendency to combine work tours with other lower-level activities for efficiency purposes and avoid travelling long distances just to shop or do a leisure activity. As expected, the utility of the least urban areas' inhabitants for choosing a shopping, leisure or drop-off only pattern decreases.

Parameters	**W**	**E**	**S**	**L**	**D**	**O**	**WFH**	**Remaining alt.**
Age 18-34	1.26	-0.79	0.58		1.77	0.39	1.96	-0.32
Age 35-64	1.33	-1.94	0.77	-0.38	2.20	0.40	2.25	-0.15
Age 65+	0.64	-2.79	0.55	-0.34	1.64		1.56	-0.57
Gender F	-0.27		0.22	-0.15			-0.37	
Car Ownership 1	0.37	0.30	0.19	0.31	0.39	0.25		0.46
Car Ownership 2+	0.51	0.31	0.14	0.36		0.23		0.50
HH Size M	-0.23		-0.12			-0.13		-0.21
HH Size L	-0.36		-0.44	-0.36	1.22	-0.51	-0.41	-0.30
Income M	0.34	0.40	0.27	0.49	0.56	0.31		0.46
Income H	0.23	0.34	0.23	0.66	0.46	0.31	0.71	0.52
Work Full Time	2.50			0.40	-0.41		3.12	1.11
Work Part time	2.36	1.26	0.38	0.65	-0.37		2.51	1.29
Student	0.68	2.57		0.46	-0.34		0.78	0.76
Urban Level 2	0.14				0.29		0.17	0.16
Urban Level 3	0.12				0.32			0.21
Urban Level 4	0.20			-0.16	-0.72			0.21
Urban Level 5	0.22		-0.25	-0.26	-0.33		0.36	

Figure 5.5: Estimated parameters for the first stage model with alternative-specific parameters

This model specification improved on the previous restricted first-stage model by increasing the $\bar{\rho}^2$ from 0.238 to 0.291 and significantly improving the log-likelihood of the model. The normalized absolute error for the complete model (first stage and second stage combined) on the 20% validation dataset decreased from 5% to 3%, indicating a better fit. Furthermore, as can be seen from fig. 5.4, the new model specification significantly improved the performance within subgroups (green bars much closer to the observed orange bars) by distinguishing the effects of attributes across different single-purpose tour patterns. This demonstrates the value of capturing heterogeneous preferences across alternatives, especially for improving the behavioral realism of tour behavior in population segments. In conclusion, based on these indicators and the likelihood ratio test, the new unrestricted MNL model with alternative-specific parameters for 6 alternatives is significantly better than the final restricted NL model from section 5.1, and is accepted as the final model for the first stage.

5.4 Case study

The case study is carried out based on the method provided in Section 3.4.3. A set of 16846 individuals (20 % of final tour data from 2022-2023) from ODiN and their respective attributes are used to simulate the generation of tours using the developed 2-stage discrete choice model and the frequency-based method followed in the Food Valley project. There are in total 112 trip chains that are modeled and compared from both models. The predicted counts for the 20 most observed trip chains that are modelled from both methods are compared with the observed counts in Fig. 5.6. The plot shows that the developed DCM predicts with high accuracy the observed trip chains, slightly over- or under-estimating some of them, but closer to reality compared to the frequency-based method. This is more obvious for the less occurring trip chains (more than one activity), which the frequency-based method consistently overestimates. The normalized absolute error (NAE) of the discrete choice model is also much lower (3%) than the other method (20%), indicating a strong performance of the DCM in predicting trip-chain frequencies of different population segments.

The DCM outperforms the frequency-based method mainly due to its high flexibility and robustness. The frequency-based method appears to overestimate the occurrence of longer trip chains that include two or more activities. This bias is likely due to the way that method handles the long tours by intensively shortening them and increasing the persona weights of the shortened tours to account for that. This adjustment can inflate the weights of longer, less frequently observed tours, leading to inaccurate estimations. On the other hand, the discrete choice model developed in this research does not fully rely on weights for estimating the tour rates, but instead uses a behavioral framework to estimate probabilities based on observed effects of attributes. Even though the discrete choice model also adjusts the weights of some shortened tours (see section 3.3.3), this process is much less intensive and has a reduced effect compared to the frequency-based method to avoid introducing bias in the model. As a result, it provides a more accurate reflection of real-world behavior, especially for complex trip chains.

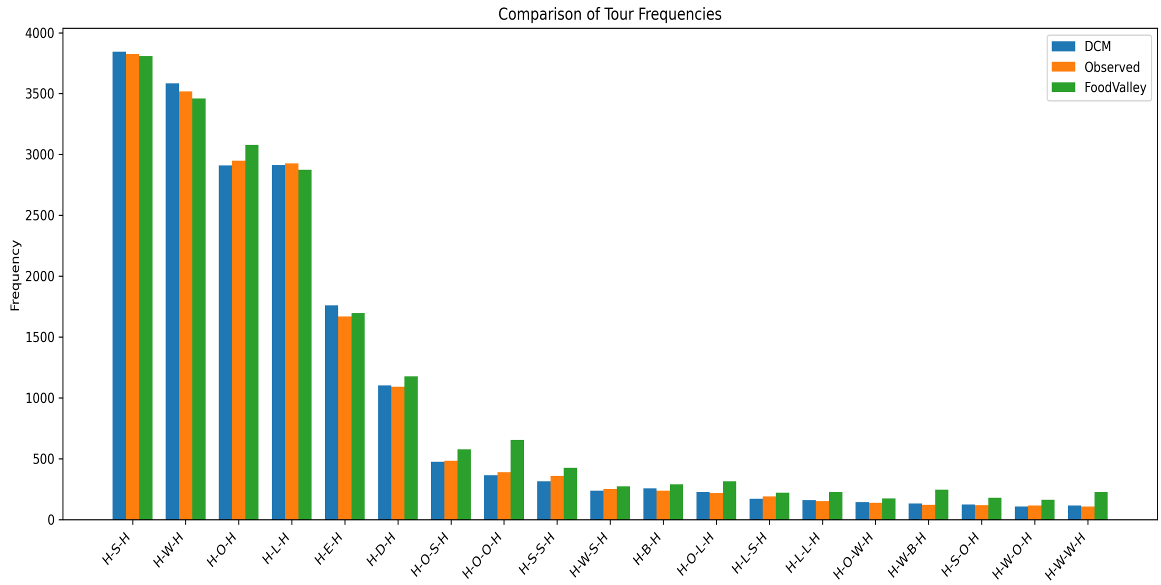


Figure 5.6: Observed and estimated counts of trip chains modeled from both methods

5.5 Results Conclusion

The estimation results for the first-stage pattern model and second-stage primary purpose models revealed different sets of attributes affecting pattern and trip-chain choice for different types of primary purposes. The two-stage modeling framework provided insightful findings into individuals' daily travel behavior. In the first-stage pattern model, which classifies complete tour patterns, the

Nested Logit (NL) model (stay home or travel nests) slightly outperformed the Multinomial Logit (MNL) model for the initial specification of utility functions and parameters, as concluded by the likelihood ratio test. However, after significant discrepancies were identified during the validation of the model within subgroups (students, elderly), the first-stage model was respecified with alternative-specific parameters for single tour-type patterns. The new MNL structure with the additional parameters significantly improved over the restricted NL model by increasing the likelihood, $\bar{\rho}^2$, and the performance within subgroups. Almost all attributes for the first stage model have a significant effect on the utilities of different daily pattern alternatives, except for the zone level. The multinomial logit structure with alternative-specific parameters is chosen to model the first-stage pattern choice as the nest parameter for the respecified model was not significantly different from 1.

In the second-stage primary purpose models, the Multinomial Logit (MNL) structure was selected for all tour purposes. Although some nested specifications showed statistically significant nest coefficients, they did not provide improvements in model fit, and the likelihood ratio tests did not support the added complexity of a Nested Logit (NL) structure. Among the explanatory variables, age and occupation status appeared most consistently as significant attributes across the different models. In contrast, urbanization level was only significant for escort and other tours, while zone level variables did not appear as significant in any of the models. Overall, the models demonstrated good explanatory power, as supported by their performance in estimating the counts of different trip chains from the validation datasets, which seem to replicate well the observed tour counts.

The case study also demonstrated the high predictive power of the developed DCM compared to a current frequency-based method. The high accuracy of the predictions of different trip chains consistently outperformed the other method, especially for the long tours, mainly due to the behavioral choice foundation that the DCM offers, tackling the drawbacks of intensively shortening trip chains and fully relying on adjusted persona weights in the frequency-based method that results in overerestimation of complex trip chains.

6 Conclusion

This final chapter presents the conclusions of the study by answering the sub-research questions and the main research question. It reflects on the key findings from the analysis, discusses their implications, and provides recommendations for future research to further improve tour-based modeling and understanding of travel behavior.

The main research question aimed at identifying the model structure and predictors that provide high explanatory power for the tour generation process of a transport model is answered in this section. To support the main question, several subquestions were first defined and elaborated in different parts of this research.

1. Which rules should be used to form tours from trip data and aggregate different types of trip chains (travel purpose combinations)?

   To address the first sub-question, a structured framework was developed based on insights from the literature and characteristics of the ODiN dataset. Tours should be defined with a base location where they usually start and end to form trip chains. The home, being the main location where people depart and arrive, is used to form home-based trip chains, in line with typical daily travel behavior. Each trip reported in the ODiN dataset should be categorized into one of these purposes: Work, Business, Education, Shopping, Leisure, Escort, or Other, covering most of the motives for which people travel daily. Then, the individual trips are chained to form tours, based on the records of the surveyed individuals.

   For tours involving multiple activities, a primary purpose should be assigned based on a hierarchical priority order (Work-Education-Shop-Leisure-Escort-Other) that reflects the relative importance of different activity types in shaping travel behavior. This allows the classification of tours into behaviorally consistent categories, even when multiple stops were made, assuming that each tour with multiple activities has a main activity, around which the other activities are planned. This method ensures that the combinatorial complexity of the trip chains is reduced into categories for aggregated data analysis and behaviorally grounded model development.

2. Which methodological approach is most suitable for estimating the frequencies of different tour types taken by specific population segments?

   Literature review about existing methodological frameworks for estimating tour frequencies in Section 2.3 revealed a few methods that are suitable for this purpose: classic regression methods, machine learning regression, and discrete choice models. After comparing these approaches, it was concluded that Discrete Choice Modeling (DCM) provides the most appropriate framework for the required model. Although other approaches like machine learning and regression techniques were taken into consideration, DCMs are notable for their abilities to represent decision-making processes at the individual level and their foundation in behavioral theory. By associating utilities with different choice alternatives, they provide a solid probabilistic and interpretable representation of travel behavior.

   In contrast to black-box machine learning models, DCMs allow for transparency in parameter interpretation, offering valuable insights into how socio-demographic and spatial characteristics affect tour generation. Furthermore, compared to regression models, they provide the possibility to explicitly specify different choice structures and relationships between alternatives using nests to group choice options sharing similarities. Specifically, multinomial logit (MNL) and nested logit (NL) structures can be used to develop model alternatives for tour generation. These advantages make DCMs well-suited not only for prediction but also for understanding the underlying behavioral mechanisms that drive tour generation. For these reasons, DCMs were selected as the modeling approach in this research.

3. Which explanatory attributes available in the data can be used to supply the model development?

   A set of explanatory attributes is needed to form the utility functions of the alternatives that explain the choices made by travelers. This study combined insights from the literature with an exploratory analysis of the ODiN travel survey. A wide range of personal and zonal attributes were considered, many of which are well-documented in previous research as important predictors for tour generation. These include age, gender, household income, household size, occupation, car ownership, and urban level. All of these attributes are available in the ODiN dataset and show meaningful variation in tour rates across different categories of primary purposes (Section 4.4), supporting their relevance for inclusion in the model.

   The only attribute considered that was not directly available in ODiN is the zone level, which was developed in a separate study aiming to represent residential accessibility of postal code areas using a combined index of household and public transport density. Overall, the explanatory attributes chosen provide a strong behavioral and empirical foundation for modeling tour generation. For an overview of the selected attributes and their respective categories, see Table 4.2.

4. Which performance metrics should be utilized to measure the models' abilities to predict tour frequencies?

   The performance evaluation of the developed models makes use of a combination of metrics from discrete choice modeling and validation strategies to compare the developed model structures and ensure their validity. The $\bar{\rho^2}$ (Rho-bar-squared) metric is used as a goodness-of-fit indicator that compares the likelihood of the model with the selected parameters to a null model where all parameters are 0. The $\bar{\rho^2}$ improves over the classic ρ^2 as it adjusts for the number of parameters, making it more appropriate when comparing models with varying complexity.

   In addition, predictive validation is performed by splitting the data into training (80%) and validation (20%) subsets. Model performance is then evaluated by comparing predicted and observed tour frequencies on the validation data, using choice probabilities derived from utility-based formulas for Logit models. Furthermore, an out-of-sample validation is conducted using data from different years of the ODiN survey. This allows testing the generalizability of the model across time. For a more intuitive evaluation, the total normalized absolute error is used to measure aggregate model accuracy, representing the proportion of total error ($Estimated - Observed$) relative to the observed total volume. To statistically assess model improvements across different structures, the likelihood ratio test is applied. This test determines whether a more complex (unrestricted) model significantly outperforms a simpler (restricted) one by comparing their final log-likelihoods.

   The practical performance of the final developed tour generation model is evaluated in a case study by comparing it with the existing tour frequency model used by Haskoning in the Food Valley project (see Section 3.4.3). Both models are trained on the same ODiN dataset, and their predictive capabilities are evaluated on an unseen test population to ensure unbiased comparison. The evaluation focuses on comparing how well each model replicates observed trip chain frequencies in the validation data. A complete set of tours is generated from both models, and only the trip chains that are found in both approaches are included in the comparison. This case study provides insights into whether the new model offers improvements in capturing real-world travel behavior.

The answers to the subquestions paved the way to construct a robust methodology for answering the main question: **Which model structure and predictors should be used to predict the frequencies of various tour types across population segments for tour generation?**

The number of different trip chains found in the data is very large (many alternatives), so based on the principles from Bowman and Ben-Akiva (2001), the first part of the model chooses a daily pattern with combinations of the primary purpose categories (6) that were previously defined or staying at home (first stage). Furthermore, to model the choice between different trip chains, a second-stage model should be used for each of the primary purpose categories, choosing a trip chain (or multiple) for the primary purposes that were chosen in the previous stage. The results of both model structures for the first stage revealed that most of the attributes have a significant effect and should be used for choosing a daily pattern, except for the zone level (see Fig. 5.5). The multinomial logit structure with alternative-specific parameters for 6 single tour-type patterns (W, S, E, L, D, O) and working-from-home significantly improved the likelihood of the restricted nested logit structure with a single set of parameters for all alternatives. Therefore, the MNL with alternative specific parameters replicated very well the observed choices with a normalized absolute error as low as 3% and reduced the discrepancies within subgroups, being chosen as the final model for this stage. The second stage primary purpose models were also estimated with a MNL and a NL structure with two nests (1: alternatives involving one tour and 2: alternatives with 2+ tours). Each primary purpose model has a different set of significant parameters for the selected attributes (see Fig. 5.2. The zone level attribute representing public transport accessibility and population density did not significantly contribute to the choice of trip chain in any of the purposes. The multinomial model is the final structure for all the second-stage models, as the nested models did not improve the likelihood of any of them, even though for a few of them (education, escort, other) the nest parameter was significantly higher than 1.

6.1 Discussion

The results present important findings about the effectiveness of the developed model. However, there are several implications related to data sources and methodological choices. First of all, the quality of the model heavily relies on the quality of the collected travel survey data (ODiN). The survey is designed to be as representative as possible of the population, but a significant amount of data, such as weekend records or individuals missing certain attributes (e.g. income), are excluded from the datasets, increasing the risk of bias and underrepresentation of certain groups. Furthermore, the formation of tours is carried out based on trip reports, assuming that all trips within a day were reported, the departure point of the next trip is always the arrival point of the last trip, and enforcing everyone to return home as the last location if that is not reported. While these assumptions are necessary for the modeling process, they can affect the validity of the model, especially if trips are significantly underreported, pointing out the need for more robust travel data in the future, such as GPS tracking.

An important methodological choice in this study was to categorize trip chains based on their primary activity, typically defined by the most important stop in the tour (e.g., work, education), assuming that it constrains the other less important activities. While this approach offers a practical way to reduce complexity (number of alternatives) in the upper-level pattern choice, it also introduces limitations that affect how well the model reflects real-world behavior. For example, tours often involve multiple activities, and assigning a single label risks oversimplifying the characteristics of the trip chain. This could lead to biases, particularly for tour patterns where non-primary activities can be important (like picking up children, play a significant role, even though they are considered secondary in a tour with work or education activities for example).

The prediction capabilities of the model from the validation results seem quite high, as the model replicates the observed tour patterns with high accuracy. The respecification of the first-stage model with alternative specific parameters for single tour patterns, significantly improved the performance on the specific groups (such as students or the elderly). This made possible the distinction of the effects of attributes between different tour types. However, the patterns with two or more tour types still share the non-specific parameters, so the attributes have the same effect on all the remaining patterns (other than W, E, S, L, D, O single-tour patterns). This might still cause undesirable effects, but making parameters specific to every alternative over-specifies the model, increasing the risk of overfitting.

As the developed model is trained on recent Dutch data (2022–2023), its validity has some temporal and geographical limitations. Travel behavior is known to vary significantly across regions and countries, due to differences in spatial structure, transport infrastructure, or cultural differences. Therefore, applying this model in geographical contexts that differ significantly compared to the Netherlands would require re-estimation using data that represents local behavior to ensure its validity. Similarly, temporal shifts in travel patterns were clearly observed in the data: behavior in years before 2021 is very different from that in later years, mainly due to the effects of the pandemic. The model, trained on post-2021 data, did not perform as well on earlier data as in the current context, suggesting that travel patterns have evolved. This points out the importance of periodic model re-estimation to maintain validity over time, especially in response to major societal changes (economy or policy).

6.2 Recommendations

Based on the findings and limitations identified in this research, several recommendations can be made to guide future work in improving the accuracy and robustness of tour-based travel demand modeling. The process of forming tours from survey trip data could be improved by carrying out additional validation checks that ensure the validity of the formed tours. Making sure that continuity really exists between trip departure and arrival locations and verifying the completeness of reported trips are crucial for building a dataset that reflects real behavior. Furthermore, future data collection methods could explore using more accurate data sources, such as GPS tracking, to construct movement patterns.

This study focuses only on home-based tours, assuming that all tours start and end at home. However, in the literature, a few studies model work-based tours, related to a common behavior of carrying out activities from the workplace and returning there again. These tours are treated as long trip-chains in this thesis that start and end at home, but future research could consider distinguishing between home-based and non-home-based tours, as the factors affecting those types of tours could differ.

In this research, trip chains are modeled explicitly as alternatives in the discrete choice model, resulting in a large number of possible combinations and excluding some rare patterns. Future research could explore different modeling approaches, such as first modeling the primary activity of the tour and then adding secondary activities sequentially to model tours. This would allow the model to construct realistic trip chains while reducing the number of predefined alternatives.

Discrete choice models offer the advantage of modeling different steps of the decision-making process explicitly. This thesis modeled the tour generation process in two stages, first choosing a pattern with the main activities of the day (primary purposes) and then choosing the trip chains, including secondary activities. This method tries to replicate the decision-making process that an individual goes through while deciding on daily travel. However, there can be different choice structures that can be used to describe the generation of activities and formation of tours. Therefore, exploring additional decision-making structures can help enrich insights on advantages and disadvantages of the different approaches.

Finally, the current validation approach relies on summing choice probabilities to estimate tour frequencies for each alternative that is modeled. Although effective for large samples, this method may not reflect the variability present when simulating patterns from calculated probabilities for smaller groups. An alternative to this could be applying Monte-Carlo simulation methods to generate discrete choices from probability distributions, providing a more realistic evaluation of predictive performance in future research. A deterministic approach like the frequency-based method that was presented in the case study, might produce more robust results in those cases when variability is not preferred.

A Appendix - Filtered Alternatives

Table A.1: Filtered alternatives for each primary purpose model and counts from data

Work	
Pattern	Count
H-W-H	16378
H-W-S-H	1175
H-B-H	1065
H-O-W-H	659
H-W-B-H	620
H-W-O-H	546
H-W-W-H	535
H-W-L-H	469
H-W-H&H-W-H	378
H-D-W-H	361
H-W-B-B-H	349
H-W-D-H	256
H-B-B-H	247
H-D-W-D-H	236
H-W-W-W-H	177
H-S-W-H	175
H-L-W-H	132
H-B-B-B-H	115
H-W-B-W-H	114
H-O-B-H	81
H-B-S-H	79
H-W-S-S-H	71
H-O-W-S-H	66
H-E-W-H	62
H-W-O-S-H	61
H-W-S-W-H	56
H-W-W-B-H	56
H-W-B-B-B-H	55
H-W-B-S-H	53
H-O-O-W-H	50
H-W-B-B-W-H	48
H-W-L-L-H	46
H-O-W-B-H	46
H-O-W-O-H	46
H-B-O-H	42
H-W-W-S-H	37
H-W-S-O-H	37
H-B-L-H	37
H-S-W-S-H	34
H-D-W-S-H	34
H-W-O-L-H	33
H-W-O-O-H	33
H-D-B-H	32
H-W-E-H	31
H-B-W-H	31
H-B-H&H-B-H	30
H-W-W-W-W-H	28

Shop	
Pattern	Count
H-S-H	14091
H-O-S-H	1928
H-S-H&H-S-H	1725
H-S-S-H	1305
H-L-S-H	777
H-S-O-H	508
H-D-S-H	337
H-S-L-H	312
H-O-O-S-H	227
H-S-S-S-H	181
H-S-H&H-S-H&H-S-H	168
H-O-S-O-H	167
H-O-S-S-H	158
H-S-S-H&H-S-H	158
H-S-H&H-S-S-H	154
H-O-S-H&H-S-H	142
H-S-D-H	133
H-S-H&H-O-S-H	114
H-L-S-H&H-S-H	72
H-O-L-S-H	61
H-L-S-S-H	58
H-S-H&H-S-O-H	53
H-D-S-D-H	52
H-S-O-S-H	51
H-D-S-H&H-S-H	51
H-S-H&H-L-S-H	50
H-S-O-O-H	50
H-O-S-L-H	49
H-S-S-O-H	49
H-O-S-H&H-O-S-H	45
H-L-S-L-H	43
H-S-O-H&H-S-H	41
H-S-L-S-H	37
H-L-S-O-H	35
H-S-H&H-S-L-H	29
H-D-D-S-H	29
H-L-O-S-H	28
H-D-S-S-H	28
H-S-O-L-H	27
H-L-L-S-H	26
H-S-H&H-D-S-H	26
H-S-L-H&H-S-H	26

Leisure	
Pattern	Count
H-L-H	12684
H-O-L-H	1066
H-L-H&H-L-H	852
H-L-L-H	697
H-L-O-H	459
H-D-L-H	147
H-O-O-L-H	138
H-L-D-H	135
H-O-L-O-H	131
H-L-L-L-H	92
H-D-L-D-H	85
H-O-L-L-H	60
H-L-H&H-L-L-H	50
H-O-L-H&H-L-H	49
H-L-H&H-O-L-H	44
H-L-H&H-L-H&H-L-H	42
H-L-O-O-H	40
H-L-L-O-H	30
H-L-O-L-H	29
H-L-H&H-L-O-H	22
H-L-O-H&H-L-H	21
H-L-L-H&H-L-H	19

Education	
Pattern	Count
H-E-H	7920
H-E-S-H	362
H-E-O-H	349
H-E-H&H-E-H	326
H-E-E-H	319
H-E-L-H	298
H-O-E-H	95
H-L-E-H	60
H-E-L-E-H	50
H-S-E-H	47
H-E-D-H	44
H-E-E-E-H	41
H-E-O-L-H	37
H-D-E-H	34
H-E-S-E-H	33
H-E-L-L-H	26
H-E-E-L-H	25
H-E-S-O-H	24
H-E-L-S-H	20
H-E-S-S-H	18
H-E-O-O-H	16
H-E-H&H-E-O-H	15

Escort	
Pattern	Count
H-D-H	2945
H-D-H&H-D-H	1022
H-O-D-H	203
H-D-H&H-D-H&H-D-H	192
H-D-D-H	162
H-D-H&H-O-D-H	96
H-D-O-H	89
H-D-H&H-D-D-H	41
H-D-O-D-H	39
H-D-D-D-H	34
H-D-D-H&H-D-H	28
H-O-O-D-H	27
H-O-D-H&H-D-H	22
H-D-H&H-D-O-H	16
H-D-D-H&H-D-D-H	16
H-O-D-D-H	16
H-D-D-H&H-D-H&H-D-H	15
H-O-D-O-H	13

Other	
Pattern	Count
H-O-H	13839
H-O-O-H	1842
H-O-H&H-O-H	611
H-O-O-O-H	406
H-O-O-H&H-O-H	118
H-O-O-O-O-H	96
H-O-H&H-O-O-H	70

B Appendix - Estimated parameters and results

Table B.1: Estimation results for the work primary purpose model (Multinomial and Nested)

Parameter	Alternative	MNL	NL
ASC_0	H-B-B-B-H	-4.79	-4.79
ASC_1	H-B-B-H	-4.11	-4.11
ASC_10	H-D-W-H	-3.75	-3.75
ASC_11	H-D-W-S-H	-6.12	-6.12
ASC_12	H-E-W-H	-5.36	-5.36
ASC_13	H-L-W-H	-4.86	-4.86
ASC_14	H-O-B-H	-5.14	-5.14
ASC_15	H-O-O-W-H	-5.83	-5.83
ASC_16	H-O-W-B-H	-5.90	-5.90
ASC_17	H-O-W-H	-3.14	-3.14
ASC_18	H-O-W-O-H	-5.74	-5.74
ASC_19	H-O-W-S-H	-5.43	-5.43
ASC_2	H-B-H	-2.60	-2.60
ASC_20	H-S-W-H	-4.62	-4.62
ASC_21	H-S-W-S-H	-6.33	-6.33
ASC_22	H-W-B-B-B-H	-5.64	-5.64
ASC_23	H-W-B-B-H	-3.78	-3.78
ASC_24	H-W-B-B-W-H	-5.75	-5.75
ASC_25	H-W-B-H	-3.22	-3.22
ASC_26	H-W-B-S-H	-5.70	-5.70
ASC_27	H-W-B-W-H	-4.85	-4.85
ASC_28	H-W-D-H	-4.14	-4.14
ASC_29	H-W-E-H	-6.10	-6.10
ASC_3	2 x H-B-H	-6.24	-5.78
ASC_31	2 x H-W-H	-3.71	-3.69
ASC_32	H-W-L-H	-3.61	-3.61
ASC_33	H-W-L-L-H	-6.01	-6.01
ASC_34	H-W-O-H	-3.41	-3.41
ASC_35	H-W-O-L-H	-6.20	-6.20
ASC_36	H-W-O-O-H	-6.20	-6.20
ASC_37	H-W-O-S-H	-5.53	-5.53
ASC_38	H-W-S-H	-2.62	-2.62
ASC_39	H-W-S-O-H	-6.03	-6.03
ASC_4	H-B-L-H	-5.89	-5.89
ASC_40	H-W-S-S-H	-5.26	-5.26
ASC_41	H-W-S-W-H	-5.80	-5.80
ASC_42	H-W-W-B-H	-5.65	-5.65
ASC_43	H-W-W-H	-3.35	-3.35
ASC_44	H-W-W-S-H	-5.68	-5.68
ASC_45	H-W-W-W-H	-4.51	-4.51
ASC_46	H-W-W-W-W-H	-6.24	-6.24
ASC_5	H-B-O-H	-5.79	-5.79
ASC_6	H-B-S-H	-5.32	-5.32
ASC_7	H-B-W-H	-6.36	-6.36
ASC_8	H-D-B-H	-5.91	-5.91
ASC_9	H-D-W-D-H	-4.17	-4.17

Parameter	Alternative	MNL	NL
B_Age_35_64	All	0.20	0.20
B_HHSize_M	All	-0.10	-0.10
B_Income_H	All	0.08	0.08
B_Work_FT	All	-0.14	-0.14
B_Work_St	All	-0.48	-0.48
B_ZoneLevel_3	All	-0.10	-0.10
MU_two_tour	2+ tour alternatives		1.21
$\bar{\rho}^2$		0.53	0.53
Log-likelihood		-36553	-36553
Norm. abs. error		0.057	0.057

Table B.2: Estimation results for the shop primary purpose model (Multinomial and Nested)

Parameter	Alternative	MNL	NL
ASC_0	H-D-D-S-H	-6.59	-6.54
ASC_1	H-D-S-D-H	-5.81	-5.80
ASC_10	H-L-S-O-H	-6.41	-6.38
ASC_11	H-L-S-S-H	-5.92	-5.91
ASC_12	H-O-L-S-H	-5.67	-5.66
ASC_13	H-O-O-S-H	-4.42	-4.41
ASC_14	H-O-S-H	-2.24	-2.23
ASC_15	2 x H-O-S-H	-5.90	-5.89
ASC_16	H-O-S-H & H-S-H	-4.90	-4.89
ASC_17	H-O-S-L-H	-5.92	-5.91
ASC_18	H-O-S-O-H	-4.57	-4.56
ASC_19	H-O-S-S-H	-4.68	-4.67
ASC_2	H-D-S-H	-3.98	-3.97
ASC_20	H-S-D-H	-4.89	-4.88
ASC_22	H-S-H & H-D-S-H	-6.53	-6.52
ASC_23	H-S-H & H-L-S-H	-6.04	-6.03
ASC_24	H-S-H & H-O-S-H	-5.15	-5.14
ASC_25	2 x H-S-H	-2.41	-2.40
ASC_26	3 x H-S-H	-4.90	-4.89
ASC_27	H-S-H & H-S-L-S	-6.36	-6.35
ASC_28	H-S-H & H-S-O-H	-5.82	-5.81
ASC_29	H-S-H & H-S-S-H	-4.77	-4.76
ASC_3	H-D-S-H & H-S-H	-5.92	-5.91
ASC_30	H-S-L-H	-4.11	-4.10
ASC_31	H-S-L-H & H-S-H	-6.76	-6.76
ASC_32	H-S-L-S-H	-6.11	-6.10
ASC_33	H-S-O-H	-3.60	-3.59
ASC_34	H-S-O-H & H-S-H	-5.99	-5.98
ASC_35	H-S-O-L-H	-6.66	-6.60
ASC_36	H-S-O-O-H	-5.85	-5.84
ASC_37	H-S-O-S-H	-5.97	-5.96
ASC_38	H-S-S-H	-2.70	-2.69
ASC_39	H-S-S-H & H-S-H	-4.91	-4.90
ASC_4	H-D-S-S-H	-6.66	-6.61
ASC_40	H-S-S-O-H	-6.14	-6.13
ASC_41	H-S-S-S-H	-4.63	-4.62
ASC_5	H-L-L-S-H	-6.74	-6.68
ASC_6	H-L-O-S-H	-6.28	-6.26
ASC_7	H-L-S-H	-3.24	-3.23
ASC_8	H-L-S-H & H-S-H	-5.73	-5.72
ASC_9	H-L-S-L-H	-6.02	-6.01
B_Age_18_34	All	0.13	0.12
B_Age_35_64	All	0.21	0.21
B_CarOwnership_1	All	0.26	0.25
B_CarOwnership_2	All	0.23	0.23
B_Gender_F	All	0.15	0.14
B_Work_FT	All	-0.21	-0.21
B_Work_PT	All	-0.15	-0.15
B_Work_St	All	-0.31	-0.31
MU_two_tour	2+ tour alternatives		1.00
$\bar{\rho}^2$		0.522	0.522
Log-likelihood	51	-33664	-33664
Norm. abs. error		0.05	0.05

Table B.3: Estimation results for the leisure primary purpose model (Multinomial and Nested)

Parameter	**Alternative**	**MNL**	**NL**
ASC_0	H-D-L-D-H	-4.57	-4.57
ASC_1	H-D-L-H	-4.12	-4.12
ASC_10	H-L-L-H & H-L-H	-6.72	-6.25
ASC_11	H-L-L-L-H	-4.55	-4.55
ASC_12	H-L-L-O-H	-5.58	-5.58
ASC_13	H-L-O-H	-3.00	-3.00
ASC_14	H-L-O-H & H-L-H	-6.18	-5.76
ASC_15	H-L-O-L-H	-5.84	-5.84
ASC_16	H-L-O-O-H	-5.30	-5.30
ASC_17	H-O-L-H	-2.21	-2.21
ASC_18	H-O-L-H & H-L-H	-5.27	-4.95
ASC_19	H-O-L-L-H	-5.07	-5.07
ASC_2	H-L-D-H	-4.24	-4.24
ASC_20	H-O-L-O-H	-4.13	-4.13
ASC_21	H-O-O-L-H	-4.20	-4.20
ASC_4	2 X H-L-H	-2.40	-2.37
ASC_5	3 X H-L-H	-5.26	-4.94
ASC_6	H-L-H & H-L-L-H	-5.46	-5.12
ASC_7	H-L-H & H-L-O-H	-5.93	-5.54
ASC_8	H-L-H & H-O-L-H	-5.28	-4.96
ASC_9	H-L-L-H	-2.54	-2.54
B_Age_35_64	All	0.12	0.12
B_CarOwnership_1	All	-0.11	-0.11
B_Work_FT	All	-0.41	-0.41
B_Work_PT	All	-0.37	-0.37
B_Work_St	All	-0.34	-0.34
MU_two_tour	2+ tour alternatives		1.11
$\bar{\rho}^2$		0.627	0.627
Log-likelihood		-15262	-15262
Norm. abs. error		0.048	0.048

Table B.4: Estimation results for the education primary purpose model (Multinomial and Nested)

Parameter	**Alternative**	**MNL**	**NL**
ASC_0	H-D-E-H	-5.07	-5.07
ASC_1	H-E-D-H	-4.67	-4.67
ASC_10	H-E-L-L-H	-5.40	-5.40
ASC_11	H-E-L-S-H	-5.90	-5.89
ASC_12	H-E-O-H	-2.86	-2.86
ASC_13	H-E-O-L-H	-5.13	-5.13
ASC_14	H-E-O-O-H	-6.17	-6.15
ASC_15	H-E-S-E-H	-5.59	-5.59
ASC_16	H-E-S-H	-2.88	-2.88
ASC_17	H-E-S-O-H	-5.40	-5.40
ASC_18	H-E-S-S-H	-5.78	-5.77
ASC_19	H-L-E-H	-4.72	-4.72
ASC_2	H-E-E-E-H	-4.80	-4.80
ASC_20	H-O-E-H	-4.09	-4.09
ASC_21	H-S-E-H	-4.84	-4.84
ASC_3	H-E-E-H	-3.02	-3.02
ASC_4	H-E-E-L-H	-5.72	-5.71
ASC_6	2 X H-E-H	-2.92	-2.90
ASC_7	H-E-H & H-E-O-H	-6.26	-4.05
ASC_8	H-E-L-E-H	-4.74	-4.74
ASC_9	H-E-L-H	-3.00	-3.00
B_Age_35_64	All	0.31	0.31
B_Gender_F	All	0.21	0.21
B_HHSize_L	All	-0.47	-0.47
B_HHSize_M	All	-0.43	-0.43
B_Income_H	All	0.17	0.17
MU_two_tour	2+ tour alternatives		2.89
$\bar{\rho}^2$		0.652	0.652
Log-likelihood		-8713.7	-8713.7
Norm. abs. error		0.057	0.058

Table B.5: Estimation results for the escort primary purpose model (Multinomial and Nested)

Parameter	**Alternative**	**MNL**	**NL**
ASC_0	H-D-D-D-H	-5.04	-5.04
ASC_1	H-D-D-H	-3.57	-3.57
ASC_10	H-D-H & H-O-D-H	-4.05	-2.95
ASC_11	H-D-O-D-H	-4.76	-4.76
ASC_12	H-D-O-H	-4.05	-4.05
ASC_13	H-O-D-D-H	-5.85	-5.85
ASC_14	H-O-D-H	-3.27	-3.27
ASC_15	H-O-D-H & H-D-H	-5.40	-3.76
ASC_16	H-O-D-O-H	-5.91	-5.91
ASC_17	H-O-O-D-H	-5.44	-5.44
ASC_2	H-D-D-H & H-D-D-H	-5.86	-4.03
ASC_3	H-D-D-H & H-D-H	-5.65	-3.91
ASC_4	H-D-D-H & 2X H-D-H	-6.16	-4.21
ASC_6	H-D-H & H-D-D-H	-4.91	-3.47
ASC_7	2 X H-D-H	-1.67	-1.53
ASC_8	3 X H-D-H	-3.30	-2.51
ASC_9	H-D-H & H-D-O-H	-5.62	-3.89
B_Age_35_64	All	0.16	0.16
B_Gender_F	All	0.30	0.30
B_Sted_2	All	0.45	0.45
B_Sted_3	All	0.33	0.33
B_Sted_4	All	0.49	0.49
B_Sted_5	All	0.45	0.45
MU_two_tour	2+ tour alternatives		1.68
$\bar{\rho}^2$		0.502	0.502
Log-likelihood		-5700.4	-5700.4
Norm. abs. error		0.078	0.078

Table B.6: Estimation results for the other primary purpose model (Multinomial and Nested)

Parameter	**Alternative**	**MNL**	**NL**
ASC_1	2 x H-O-H	-3.38	-3.20
ASC_2	H-O-H & H-O-O-H	-5.57	-3.82
ASC_3	H-O-O-H	-2.35	-2.35
ASC_4	H-O-O-H & H-O-H	-5.17	-3.71
ASC_5	H-O-O-O-H	-3.81	-3.81
ASC_6	H-O-O-O-O-H	-5.34	-5.34
B_Age_18_34	All	0.19	0.19
B_Age_35_64	All	0.48	0.48
B_CarOwnership_1	All	0.40	0.40
B_CarOwnership_2	All	0.46	0.46
B_HHSize_L	All	-0.36	-0.36
B_Sted_2	All	0.12	0.12
B_Work_FT	All	-0.41	-0.41
B_Work_PT	All	-0.22	-0.22
B_Work_St	All	-0.52	-0.52
MU_two_tour	2+ tour alternatives		3.53
$\bar{\rho}^2$		0.641	0.641
Log-likelihood		-9467.6	-9467.6
Norm. abs. error		0.012	0.0126

Table B.7: Estimation results for the first stage model with alternative specific parameters

Parameter	Alternative	MNL
ASC_0	D	-5.28
ASC_1	D-O	-4.63
ASC_10	L	-1.20
ASC_11	L-D	-4.52
ASC_12	L-O	-3.67
ASC_13	O	-1.19
ASC_14	S	-0.91
ASC_15	S-D	-3.43
ASC_16	S-D-O	-5.03
ASC_17	S-L	-2.54
ASC_18	S-L-D	-4.74
ASC_19	S-L-O	-4.25
ASC_2	E	-1.81
ASC_20	S-O	-2.14
ASC_21	W	-2.74
ASC_22	W-D	-3.87
ASC_23	W-D-O	-5.65
ASC_24	W-E	-4.46
ASC_25	W-L	-2.38
ASC_26	W-L-D	-5.56
ASC_27	W-L-O	-5.02
ASC_28	W-O	-2.30
ASC_29	W-S	-2.62
ASC_3	E-D	-5.87
ASC_30	W-S-D	-5.33
ASC_31	W-S-L	-4.68
ASC_32	W-S-O	-4.27
ASC_33	WFH	-2.98
ASC_4	E-L	-2.46
ASC_5	E-L-O	-5.59
ASC_6	E-O	-3.77
ASC_7	E-S	-3.81
ASC_8	E-S-L	-5.15
B_Age_18_34	Remaining	-0.32
B_Age_18_34_D	D	1.77
B_Age_18_34_E	E	-0.79
B_Age_18_34_O	O	0.39
B_Age_18_34_S	S	0.58
B_Age_18_34_W	W	1.26
B_Age_35_64	Remaining	-0.15
B_Age_35_64_D	D	2.20
B_Age_35_64_E	E	-1.94
B_Age_35_64_L	L	-0.38
B_Age_35_64_O	O	0.40
B_Age_35_64_S	S	0.77
B_Age_35_64_W	W	1.33

Parameter	Alternative	MNL
B_Age_65	Remaining	-5.28
B_Age_65_D	D	-4.63
B_Age_65_E	E	-1.20
B_Age_65_L	L	-4.52
B_Age_65_S	S	-3.67
B_Age_65_W	W	-1.19
B_CarOwnership_1	Remaining	-0.91
B_CarOwnership_1_D	D	-3.43
B_CarOwnership_1_E	E	-5.03
B_CarOwnership_1_L	L	-2.54
B_CarOwnership_1_O	O	-4.74
B_CarOwnership_1_S	S	-4.25
B_CarOwnership_1_W	W	-1.81
B_CarOwnership_2	Remaining	-2.14
B_CarOwnership_2_E	E	-2.74
B_CarOwnership_2_L	L	-3.87
B_CarOwnership_2_O	O	-5.65
B_CarOwnership_2_S	S	-4.46
B_CarOwnership_2_W	W	-2.38
B_Gender_F_L	L	-5.56
B_Gender_F_S	S	-5.02
B_Gender_F_W	W	-2.30
B_HHSize_L	Remaining	-2.62
B_HHSize_L_D	D	-5.87
B_HHSize_L_L	L	-5.33
B_HHSize_L_O	O	-4.68
B_HHSize_L_S	S	-4.27
B_HHSize_L_W	W	-2.98
B_HHSize_M	Remaining	-2.46
B_HHSize_M_O	O	-5.59
B_HHSize_M_S	S	-3.77
B_HHSize_M_W	W	-3.81
B_Income_H	Remaining	-5.15
B_Income_H_D	D	-0.32
B_Income_H_E	E	1.77
B_Income_H_L	L	-0.79
B_Income_H_O	O	0.39
B_Income_H_S	S	0.58
B_Income_H_W	W	1.26
B_Income_M	Remaining	-0.15
B_Income_M_D	D	2.20
B_Income_M_E	E	-1.94
B_Income_M_L	L	-0.38
B_Income_M_O	O	0.40
B_Income_M_S	S	0.77
B_Income_M_W	W	1.33

Parameter	Alternative	MNL
B_Sted_2	Remaining	0.16
B_Sted_2_W	W	0.14
B_Sted_3	Remaining	0.21
B_Sted_3_W	W	0.12
B_Sted_4	Remaining	0.21
B_Sted_4_L	L	-0.16
B_Sted_4_W	W	0.20
B_Sted_5_D	D	-0.33
B_Sted_5_L	L	-0.26
B_Sted_5_S	S	-0.25
B_Sted_5_W	W	0.22
B_Work_FT	Remaining	1.11
B_Work_FT_D	D	0.29
B_Work_FT_L	L	0.40
B_Work_FT_W	W	2.50
B_Work_PT	Remaining	1.29
B_Work_PT_D	D	0.32
B_Work_PT_E	E	1.26
B_Work_PT_L	L	0.65
B_Work_PT_S	S	0.38
B_Work_PT_W	W	2.36
B_Work_St	Remaining	0.76
B_Work_St_D	D	-0.72
B_Work_St_E	E	2.57
B_Work_St_L	L	0.46
B_Work_St_W	W	0.68
$\bar{\rho}^2$		0.291
Log-likelihood		-168406
Norm. abs. error		0.048

C Appendix - Detailed results interpretation

The parameters and validation plots for each primary purpose model in the second stage are described in detail in this appendix.

C.1 Work primary purpose model

The multinomial and nested models for the work primary purpose tours with 47 alternatives are estimated, and the full set of parameters, including ASCs can be found in Table **??**. Fig. 5.3 shows the significant parameters for the personal attributes in making work tours and their effect in utility (positive in blue and negative in red). Individuals aged 35-64 derive increased utility from making a more complex work tour combined with other activities or multiple work tours on a day, compared to other age groups. This behavior might be related to the high activity participation of this age group and stopping to complete these activities on the way to/from work. Also high-income groups are more likely to make a work tour different from the simple H-W-H, possibly due to higher flexibility and financial capacity, or possible access to a car. On the other hand, individuals from a medium-sized household have a lower utility for making additional stops on work tours compared to single-person households, likely because of responsibility sharing and coordination with the partner to run errands (e.g., shopping). Working full-time and being a student (given that they make a work tour) also decrease the utility of combining work tours with other activities, as the time constraints are stricter. Urban and zone level parameters were not significantly different from 0 for explaining the trip chain choice in work tours. The MU parameter of the nest with 2 or more tours (2+) was estimated to be 1.2, which is very close to one, suggesting low correlation between the nested alternatives and modelling the choice between the nested alternatives similarly to the multinomial structure.

The estimation results in Table 5.2 show that $\bar{\rho}^2$ and log-likelihood of both model structures is the same, with a value of 0.53 for the $\bar{\rho}^2$, which is considered good a good fit as it significantly improves over the likelihood of the null model (all parameters are 0). As the nested model does not improve over the likelihood of the restricted model, it is directly rejected using the likelihood ratio test. The normalized absolute error from the two structures is also the same at around 5% as the nest parameter is very close to 1, and the choice between the nested alternatives is very similar to the multinomial structure. The estimated and observed counts of each alternative from both model structures are shown in Fig. 3.1 for the internal validation with 20% of the data not used for training (from 2022-2023), and in Fig. 3.2 for the external validation (2018-2019) dataset. The simulated choices of alternatives in the internal validation dataset reproduce well the observed counts. There is a slight underestimation of H-W-H and a slight overestimation of H-B-H, but there is no significant deviation from real-world behavior. On the other hand, testing the model on the data from 2018-2019, there appear to be significant deviations from the observed tour counts for certain alternatives, likely due to a significant shift in travel behavior after the pandemic. The work tours combined the purpose other (O) or with multiple business (B) stops seem to be substantially overestimated, and H-W-H tours underestimated (the bar for H-W-H in Fig. 3.2 is not fully shown as figure is zoomed to visualize differences in lower count alternatives). This is mainly due to more travelers combining work tours with other activities after the pandemic, rather than just doing a direct tour to work and back. As the model is trained to data from 2022-2023, it does not generalize that well over the behavior before the COVID pandemic (2019-2021). The final model structure for the work primary purpose model is the multinomial logit based on the likelihood ratio test, and no significant difference in the estimation of observed choices between the restricted and unrestricted (nested) models.

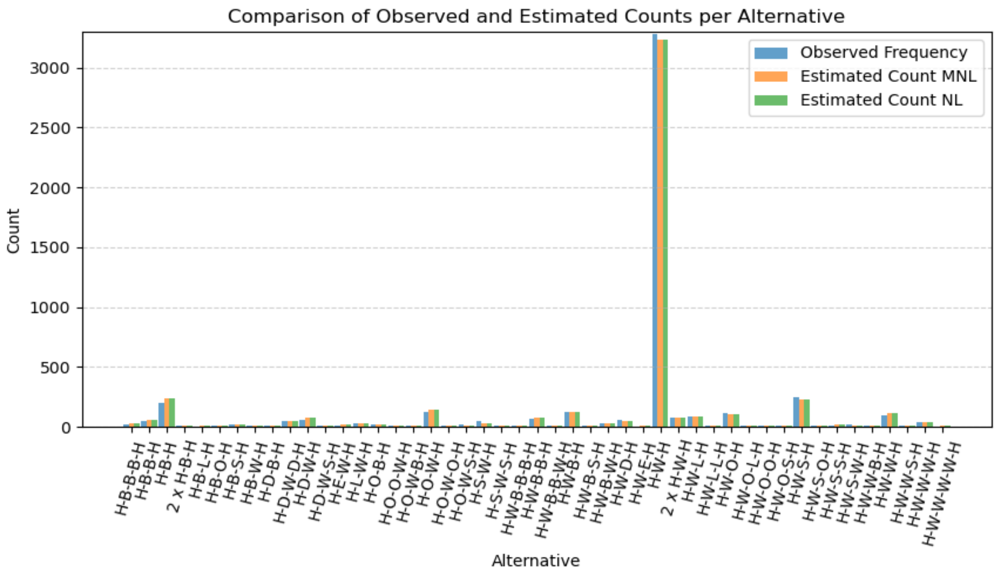


Figure 3.1: Observed and estimated counts of work primary purpose tours in the validation dataset

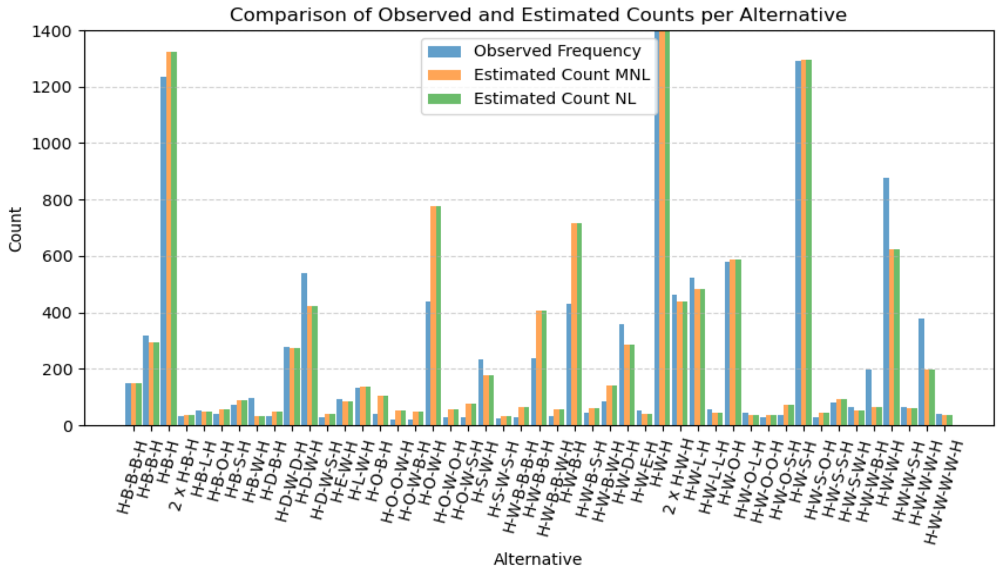


Figure 3.2: Observed and estimated counts of work primary purpose tours in the external (2018-2019) validation dataset

C.2 Education primary purpose model

The parameters of both model structures for the primary education purpose are shown in Table B.4. The age group 35-64 derives a higher utility from making an education tour combined with other lower-level activities (S, L, D, O), reflecting the high activity level of adults participating in education activities, making them chain multiple activities in one tour. Females also have a higher likelihood of making complex education tours, possibly due to their role in the household of carrying out certain tasks, such as shopping. Being in a medium or large-sized household decreases the probability of taking long education tours as the coordination and task sharing within the household increase. High-income individuals seem to be more likely to combine education tours with other activities, likely due to having more financial means to support additional activities such as leisure and shopping. The nest parameter for the 2 or more tour alternatives (2 x H-E-H and H-E-H & H-E-O-H) was estimated to be significant with a value of 2.9, revealing a strong correlation between these alternatives.

As can be seen from Table 5.2, the $\bar{\rho}^2$ of both model structures is the same with a value of 0.65, the highest of all models. The log-likelihood of the restricted and unrestricted (nested) models are also identical, pointing out that the nesting of the alternatives did not improve the model, so the nested model is rejected using the likelihood ratio test, even though the nest parameter was significantly different from 1. The normalized absolute error on the validation dataset are also almost the same (5.7% for the multinomial structure and 5.8% for the nested). The predicted counts for each alternative reproduce well the observed counts in the validation dataset, see Fig. 3.3. The external validation plot in Fig. 3.4 reveals some discrepancies in the estimation of a couple of alternatives. The alternatives with two eduction tours in a day (2 x H-E-H and H-E-H & H-E-O-H) seem to be significantly underestimated in the data from 2018-2019. That can be related to the model being estimated from data after the pandemic, when its effects might have reduced the likelihood of people doing two education tours in a day (more people study from home). Furthermore, the alternatives that combine education with other activities (H-E-L-H, H-E-O-H, H-E-S-H) are slightly overestimated, reflecting also a possible behavioral shift after the pandemic regarding increasing non-mandatory activity levels. The final model for estimating the primary education trip chains is the multinomial structure based on the likelihood ratio test and performance results.

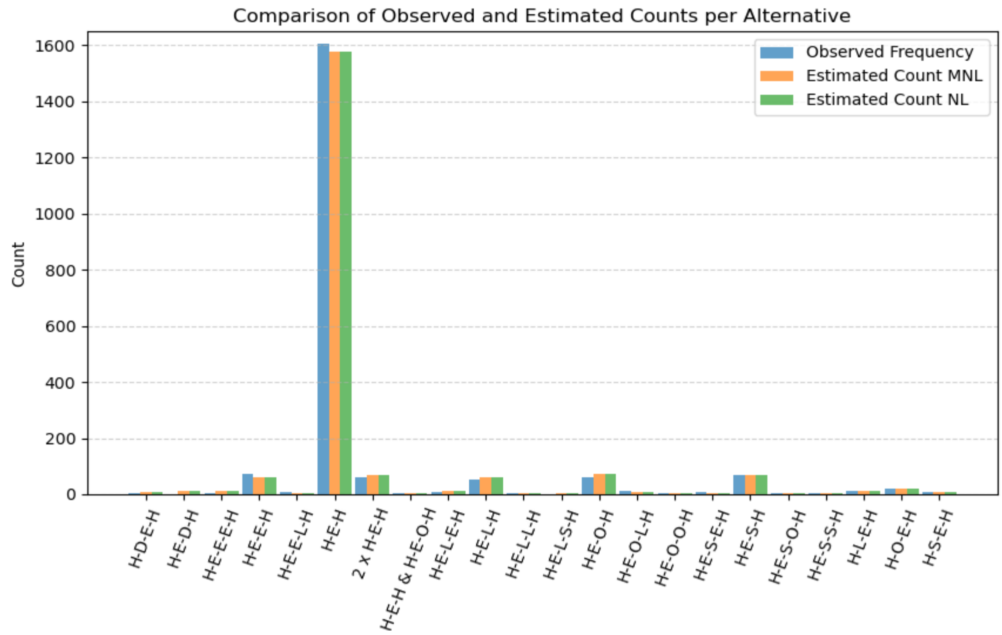


Figure 3.3: Observed and estimated counts of education primary purpose tours in the validation dataset

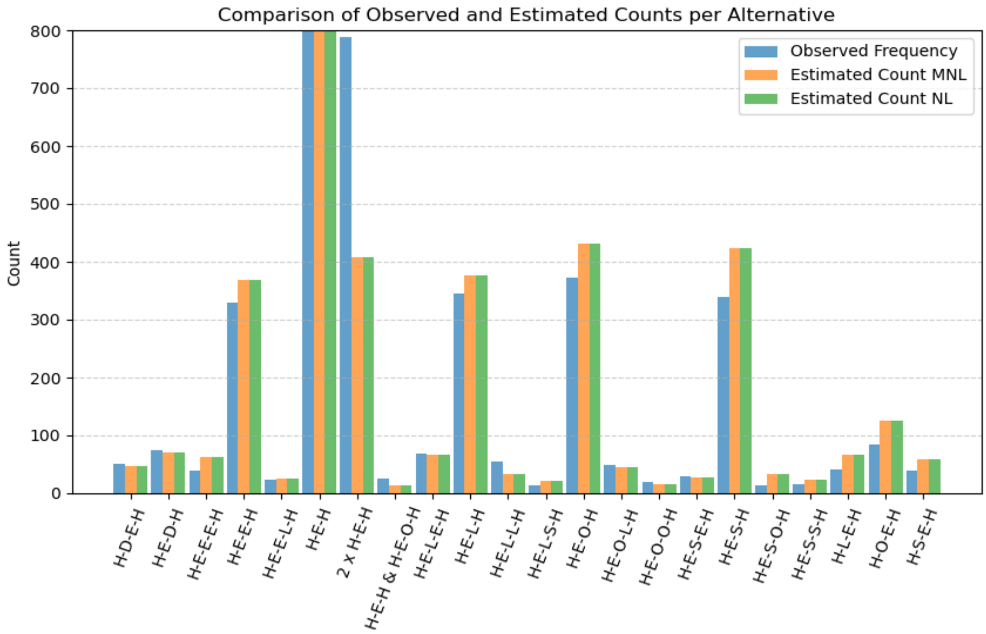


Figure 3.4: Observed and estimated counts of education primary purpose tours in the external (2018-2019) validation dataset

C.3 Shop primary purpose model

The full estimation results of both model structures for the shop primary purpose are shown in Table B.2. For this purpose, the nested structure did not output a nest parameter significantly different than 1. Therefore, the nested model is invalid, and the multinomial structure is the only and final model for the choice between shopping trip-chains. As seen from Fig. 5.2, the young (18-34) and middle-aged (35-64) adults are more likely than the other age groups to combine shopping tours with other activities (leisure, escort, other), likely due to the active lifestyle and stricter schedules for these groups, forcing them to chain multiple activities in a tour for time efficiency. Car ownership also significantly increases the likelihood of making a complex shopping tour, providing convenience (storage) and flexibility to visit multiple destinations before returning home. Females have a higher utility than males for making complex shopping tours, possibly related to their traditional role in household shopping or greater likelihood of combining errands. In contrast, having an occupation (worker or student) decreases the probability of making long shopping tours.

The model seems to explain the observed choices well with $\bar{\rho}^2$ being 0.52 and the normalized absolute error in the validation dataset around 5%. The validation plot in Fig. 3.5 shows that the model fits the observed counts of alternatives very well. As there are no significant discrepancies in the estimation of observed choices, the developed multinomial logit model for the shop primary purpose can be used to estimate the trip chains in this category.

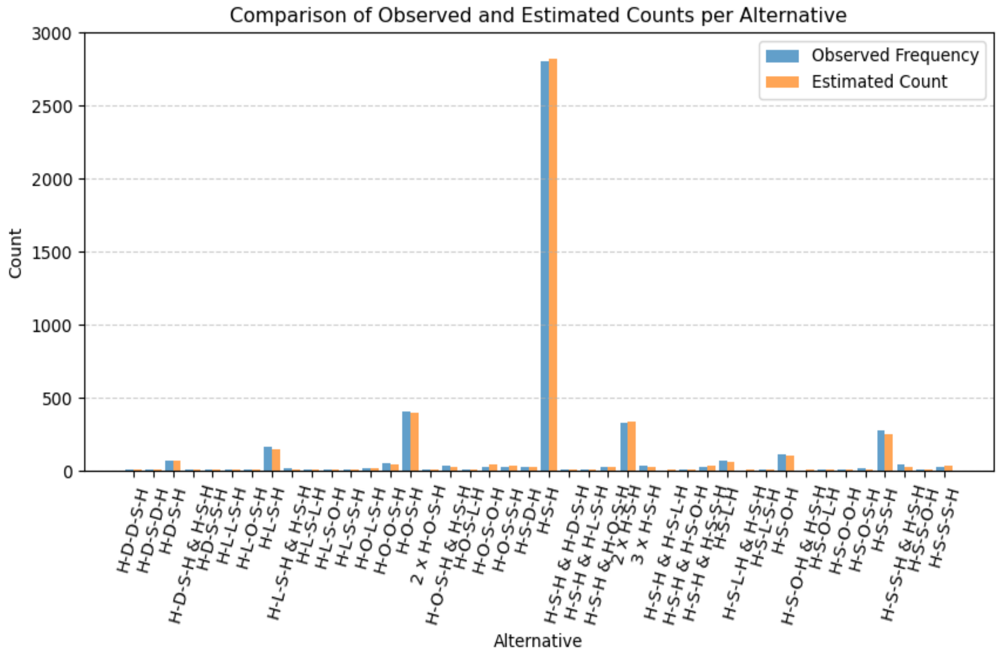


Figure 3.5: Observed and estimated counts of shop primary purpose tours in the validation dataset

C.4 Leisure primary purpose model

The model alternatives for the leisure primary purpose tours have been estimated, and the full set of parameters can be found in Table B.3. The estimated parameters for the attributes in Fig. 5.3 show that being in the age group 35-64 slightly increases the utility of making a longer leisure tour (combined with escort or/and other) compared to children, possibly due to having more responsibilities, such as dropping off or picking up kids. On the other hand, being employed or a student decreases the preference for making a longer tour compared to unemployed groups, likely because of time constraints. Car ownership also reduces the utility of making a complex leisure tour. The nest parameter of two or more tour alternatives (2+) was estimated with a value of 1.1, which suggests that the alternatives under the nest are almost independent (MU=1 means they are independent and not correlated). The other parameters were insignificant and are therefore not included in this table.

The estimation indicators in Table 5.2 show that base multinomial logit and the nested structure with two nests (1 tour or 2+ tours) yield the same $\bar{\rho}^2$ value of 0.63, which is quite high. The likelihood of the MNL is again the same as the NL structure, so the NL model is directly rejected with the likelihood ratio test, given that adding one parameter (nest) does not improve the likelihood. The normalized absolute error on the validation dataset is also the same for both model structures, as the value of MU close to 1 models the choice between the nested alternatives very similarly to the multinomial logit. The observed and estimated counts of each alternative in the validation dataset are shown in Fig. 3.6 and for the external validation in Fig.3.7. As can be seen both models provide high predictive accuracy in the validation plot. However, in the external validation plot, the model seems to be significantly overestimating leisure tours combined with the Other (O) purpose, possibly due to a change in travel behavior after 2019, indicating that travelers participate in more non-mandatory activities and combine them in longer tours.

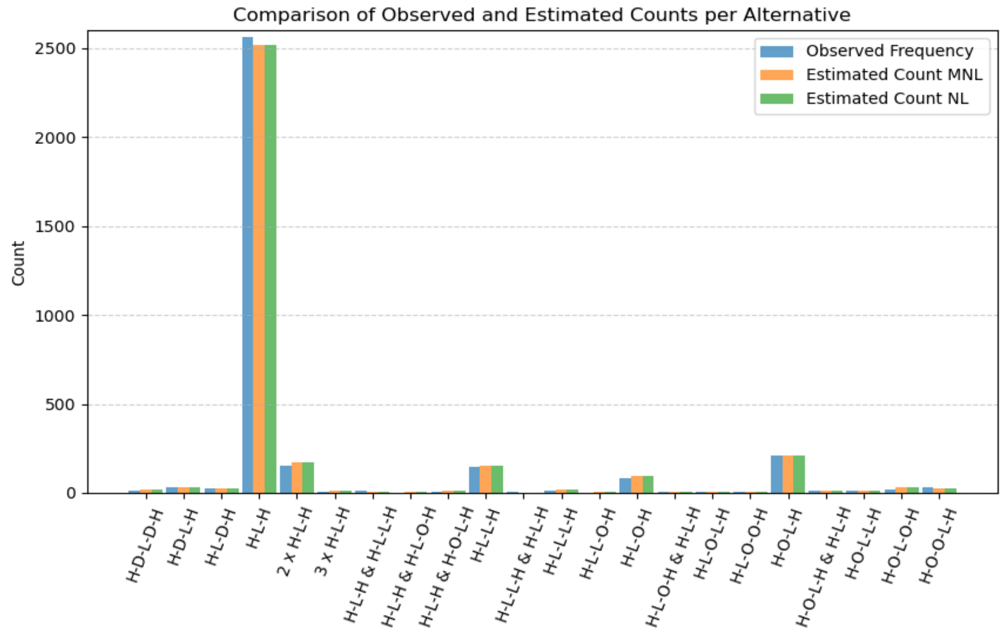


Figure 3.6: Observed and estimated counts of the multinomial logit model in the validation dataset

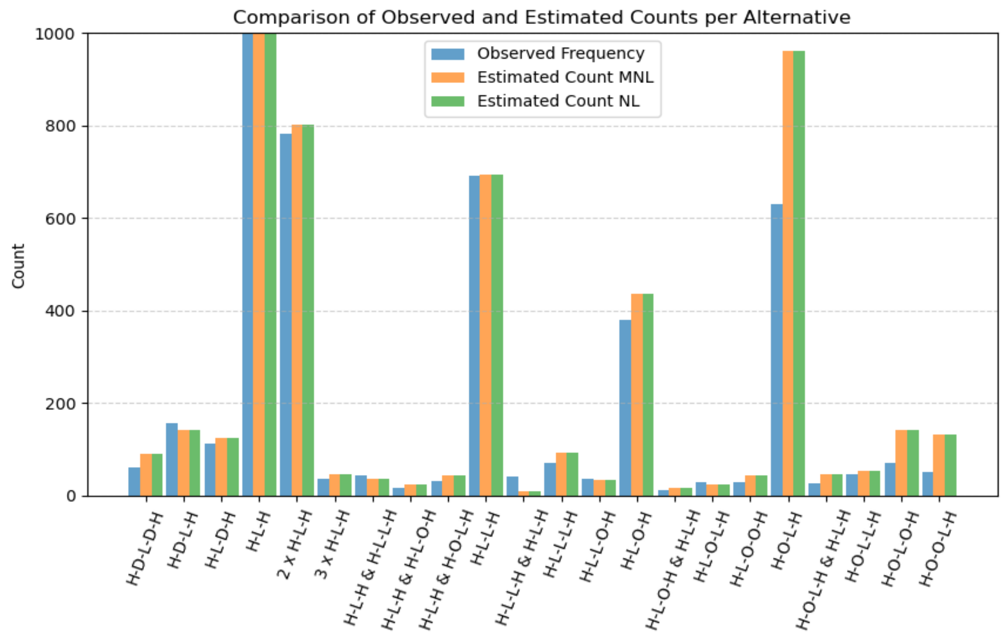


Figure 3.7: Observed and estimated counts of the nested logit model in the validation dataset

C.5 Escort primary purpose model

The estimated parameters for the escort primary purpose show that the urban level has a significant effect on choosing the trip chain for these tours. Individuals who live in less urban areas (different from urban level 1) derive a higher utility for combining escort tours with other activities, likely for cost efficiency. The multinomial and nested structure yield the same log-likelihood and normalized absolute error, leading to the choice of MNL as the final structure for the escort primary purpose model.

The validation plot in Fig. 3.8 (internal) shows that the escort primary purpose model predicts quite well the observed counts for different alternatives. There are some slight variations from the 2018-2019 external validation in Fig. 3.9 between the observed and estimated counts, but the overall fit is still quite representative of the behavior.

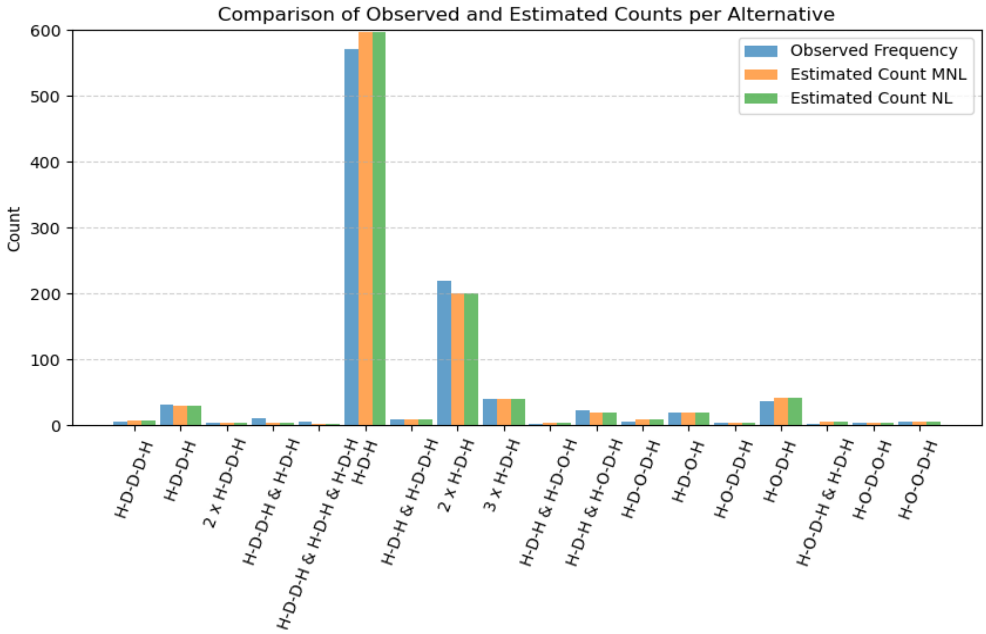


Figure 3.8: Observed and estimated counts of escort primary purpose tours in the validation dataset

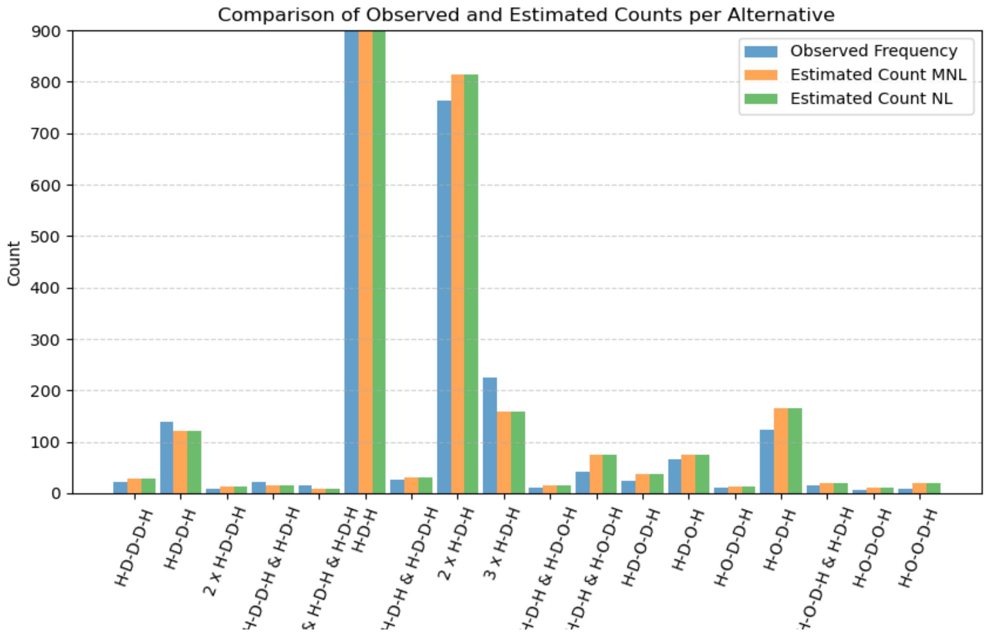


Figure 3.9: Observed and estimated counts of escort primary purpose tours in the external (2018-2019) validation dataset

C.6 Other primary purpose model

The estimation results for the other primary purpose model in Fig. 5.2 show that car ownership increases the utility of carrying out complex other tours (multiple stops with other activities), likely due to the convenience that the car offers. On the contrary, individuals with an occupation are much less likely to take long other tours, as they possibly combine these activities with work or education tours. The normalized absolute error for the other primary purpose model is the lowest of all models (1%). This can be due to the limited number of alternatives and the presence of only other (O) activities in the trip chains. The nested model, even though with a significantly different from 1 nest parameter (3.3), did not significantly improve the log-likelihood, resulting in MNL being chosen as the final structure for this model too.

The validation plot in Fig. 3.10 shows almost perfect estimations of observed counts for all alternatives, pointing out high predictive abilities for this model. The complex alternatives (other than H-O-H) are slightly overestimated in the external validation plot in Fig. 3.11, likely due to increased activity levels of non-mandatory activities after the pandemic.

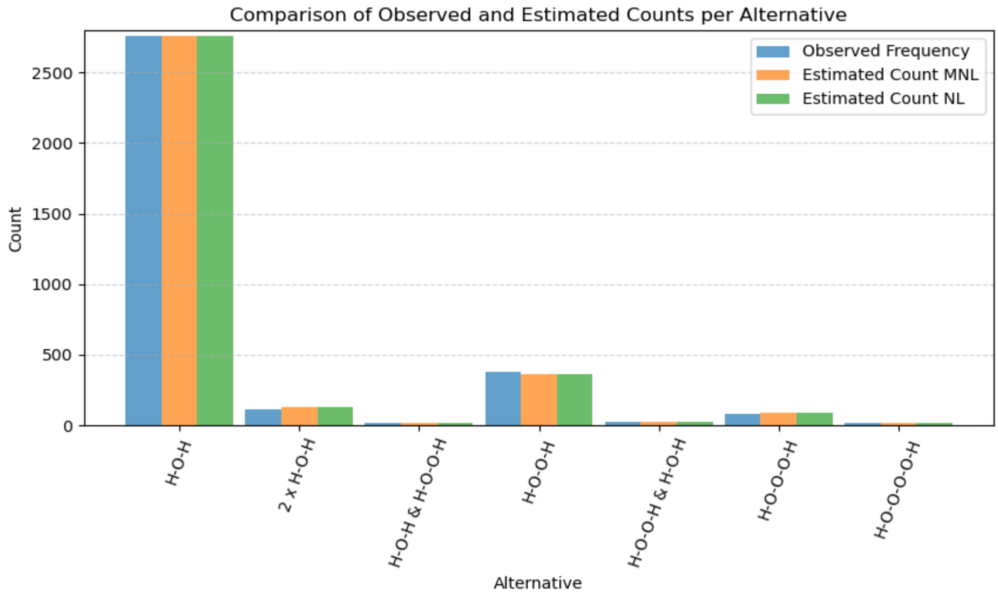


Figure 3.10: Observed and estimated counts of other primary purpose tours in the validation dataset

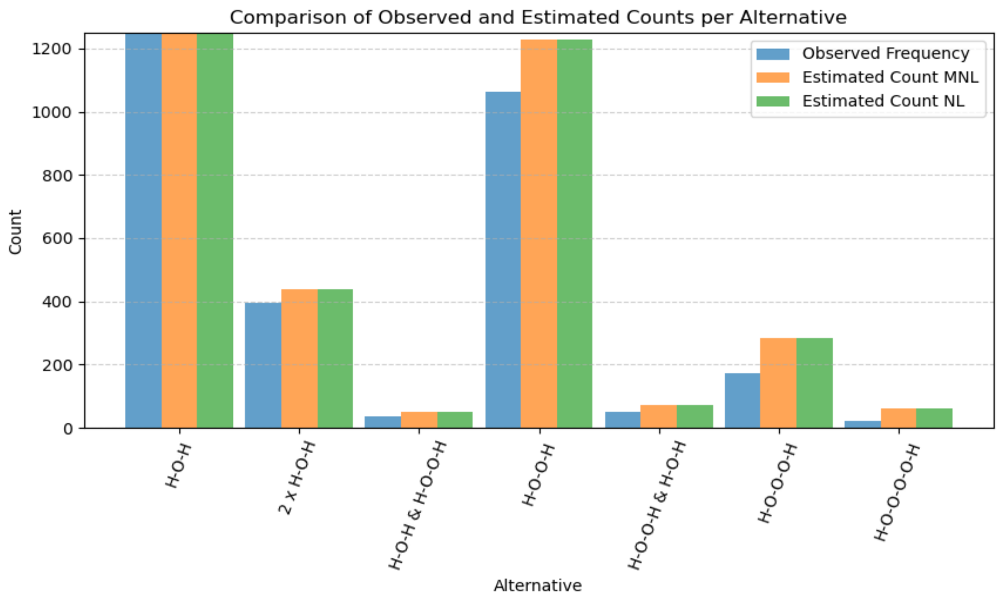


Figure 3.11: Observed and estimated counts of other primary purpose tours in the external (2018-2019) validation dataset

Bibliography

Bierlaire, M. (2016). PythonBiogeme: A Short Introduction. Technical report, Transport and Mobility Laboratory, ENAC, École Polytechnique Fédérale de Lausanne.

Bowman, J. L. and Ben-Akiva, M. E. (2001). Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A: Policy and Practice*, 35(1):1–28.

Centraal Bureau voor de Statistiek (2024). Onderweg in Nederland (ODiN).

de Jong, G., Daly, A., Pieters, M., Miller, S., Plasmeijer, R., and Hofman, F. (2007). Uncertainty in traffic forecasts: literature review and new results for The Netherlands. *Transportation*, 34(4):375–395.

de Romph, E. (2021). Persoonstypes & Tours Verkeersmodel Limburg. PLATOS 2021 – MODELLEN IN DE ACTUALITEIT.

Dumbliauskas, V. (2019). *VILNIUS GEDIMINAS TECHNICAL UNIVERSITY*.

Esmael, M., SASAKI, K., and NISHII, K. (2009). Tour-based Travel Demand Modeling using Person Trip Data and its Application to Advanced Policies. *Proceedings of the Eastern Asia Society for Transportation Studies*, 2009:112.

Festa, D. C., Condino, D., and Mazzulla, G. (2006). Experimental tour-based travel demand models. *European Journal of Operational Research*, 175(3):1472–1483.

Goel, S. and Sinha, A. K. (2008). Trip Generation Modeling using Artificial Neural Network. Technical report.

Golob, T. F. (2000). A simultaneous model of household activity participation and trip chain generation. *Transportation Research Part B: Methodological*, 34(5):355–376.

Groeneveld, Y., Delft, T., de Romph, E., Haskoning DHV Alex Nugteren, R., Haskoning DHV Adam Pel, R., and Daamen, W. (2023). Performance analysis on joint mode and destination choice: a tour-based model of Limburg. Technical report.

Guzman, L. A., Gomez, A. M., and Rivera, C. (2017). A Strategic Tour Generation Modeling within a Dynamic Land-Use and Transport Framework: A Case Study of Bogota, Colombia. *Transportation Research Procedia*, 25:2536–2551.

Hamad, K. and Obaid, L. (2022). Tour-based travel demand forecasting model for a university campus. *Transport Policy*, 117:118–137.

Heggie, I. G. and Jones, P. M. (1978). Defining domains for models of travel demand. *Transportation*, 7(2):119–125.

Ida Kristoffersson, S. B. and Algers, S. (2020). Estimation of a large-scale tour generation model taking travellers' daily tour patterns into account. *Transportation Planning and Technology*, 43(2):158–173.

Ipek Sener, A. N., Ferdous, N., and Bhat, C. R. (2009). Tour-based model development for TxDOT : evaluation and transition steps. Technical report.

Kenneth E. Train (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press, 2nd edition.

Krizek, K. J. (2003). Neighborhood services, trip purpose, and tour-based travel. *Transportation*, 30(4):387–410.

Nowrouzian, R. and Srinivasan, S. (2012). Empirical Analysis of Spatial Transferability of Tour-Generation Models. *Transportation Research Record*, 2302(1):14–22.

Ortelli, N., De Lapparent, M., and Bierlaire, M. (2023). Faster estimation of discrete choice models via weighted dataset reduction. Technical report.

Patruni, B., Fox, J., and Phillips, W. D. (2021). Model of Travel in London Phase 3: Frequency model estimation EUROPE. Technical report.

Pirra, M. and Diana, M. (2016). Classification of Tours in the U.S. National Household Travel Survey through Clustering Techniques. *Journal of Transportation Engineering*, 142(6):4016021.

Přibyl, O. and Goulias, K. G. (2005). Simulation of Daily Activity Patterns Incorporating Interactions within Households: Algorithm Overview and Performance. *Transportation Research Record*, 1926(1):135–141.

Primerano, F., Taylor, M. A. P., Pitaksringkarn, L., and Tisato, P. (2008). Defining and understanding trip chaining behaviour. *Transportation*, 35(1):55–72.

PTV Group (2024). Transportation Modelling: Challenges & Solutions.

Rossi, T. F. and Shiftan, Y. (1997). Tour Based Travel Demand Modeling in the U.S. *IFAC Proceedings Volumes*, 30(8):381–386.

Shams, K., Asgari, H., Hossan, M. S., and Jin, X. (2018). An investigation of tour generation models combining two waves of household travel surveys through pooled models. *Transportation Planning and Technology*, 41(3):229–243.

Train, K. (2002). Discrete Choice Methods with Simulation. Technical report.

Vovsha, P. (2019). Decision-Making Process Underlying Travel Behavior and Its Incorporation in Applied Travel Models. In *Advances in Intelligent Systems and Computing*, volume 805, pages 36–48. Springer Verlag.

Yagi, S. and Mohammadian, A. K. (2008). Modeling Daily Activity–Travel Tour Patterns Incorporating Activity Scheduling Decision Rules. *Transportation Research Record*, 2076(1):123–131.