Universität Zürich[UZH]

Geographisches Institut

A Leisure Attractivity Framework for Destination Choice in Travel Demand Models

GEO 511 Master's Thesis

Author: Stefano Benzoni, 20-703-278

Supervised by: Prof. Dr. Esra Suel, Antonin Danalet (antonin.danelet@sbb.ch), Davi Guggisberg (davi.guggisberg@sbb.ch)

Faculty representative: Prof. Dr. Esra Suel

24.04.2026

Universität Zürich[UZH]

Geographisches Institut

A Leisure Attractivity Framework for Destination Choice in Travel Demand Models

Stefano Benzoni

20-703-278

Master of Science in Geography

University of Zurich (UZH)

Swiss Federal Railways (SBB)

Supervisors:

Prof. Dr. Esra Suel (UZH)

Antonin Danalet (SBB)

Davi Guggisberg (SBB)

Faculty Member:

Prof. Dr. Esra Suel (UZH)

GEO 511 Master's Thesis

April 24, 2026

Acknowledgements

I would first like to express my sincere gratitude to Prof. Dr. Esra Suel for her constant support throughout this journey, for her valuable feedback, and for always providing me with interesting ideas and perspectives from which to start and develop this work.

I would also like to thank Antonin Danalet and Davi Guggisberg for their support, for their advice, and for giving me the opportunity to work in close contact with a real operational system. I am equally grateful to the SIMBA MOBi team at SBB for this important opportunity.

I would like to thank my family for the love they have always given me and for supporting me throughout my studies.

Special thanks go to Anna for her emotional support, both in good moments and in more difficult ones, for always listening to me, and for giving me good advice.

Finally, I would like to thank my friends from *P704* for making everything more enjoyable.

Abstract

Switzerland is undergoing demographic and mobility change: the population is ageing, and leisure is already the most frequent trip purpose, especially among older adults. At the same time, national transport outlooks suggest that leisure will account for an increasing share of future travel. Together, these trends strengthen the need for a better understanding and representation of leisure mobility.

To represent and predict where people choose to go for leisure, travel demand models typically rely on destination-choice formulations in which utility combines an accessibility term (or cost) with a destination-side attractivity term (or size term). The latter is often represented by a small number of variables, such as employment or resident population. While operationally convenient, such representations make it difficult to capture the diversity of leisure opportunities across places and to reveal differences in destination-side preferences across population groups and travel contexts. This thesis addresses this limitation by developing an improved representation of the attractivity term in zone-based travel demand modelling. The proposed formulation combines variables describing leisure-related supply, territorial structure, and environmental features.

As a pragmatic first step to test whether this formulation can reveal behavioural differences, the analysis distinguishes leisure destination choice along two dimensions: age group (6–64 vs. 65+) and travel context (short everyday leisure vs. long day-trip leisure), yielding four estimation segments. Within a multinomial logit discrete-choice framework, segment-specific models are estimated and validated using observed leisure destination choices from the national travel diary survey, the *Swiss Mobility and Transport Microcensus*. Model performance is assessed through out-of-sample evaluation and comparison with alternative formulations. As a proof of concept, the resulting segment-specific attractivity layer is then integrated into SIMBA MOBi, the agent-based travel demand model operated by the Swiss Federal Railways (SBB).

The results show that a richer destination-side attractivity formulation improves the representation of leisure destination choice and helps reveal heterogeneity across segments. The clearest differences emerge between short and long leisure travel contexts, while age-related differences are present but more limited. From an applied perspective, the thesis delivers a practical, flexible, and transferable framework for specifying a richer destination-side attractivity term in travel demand modelling. It also demonstrates that such a formulation can be translated into an implementation-oriented attractivity layer for operational models, while highlighting that further calibration is needed for its full benefits to emerge in a simulation environment.

Code availability. The code and materials associated with this thesis are available on GitHub at MSc_Thesis_Benzoni.

Contents

List of Figures

List of Tables

1 Introduction

This chapter first presents the empirical observations motivating the thesis, namely population ageing, the current importance of leisure in daily travel, and expected changes in future mobility patterns (Section 1.1). It then discusses why leisure mobility is particularly difficult to model (Section 1.2), formulates the specific modelling problem addressed (Section 1.3), and sets out the research objectives and questions (Section 1.4).

1.1 Swiss demographic and mobility context

This thesis is motivated by three simple empirical observations related to (i) population ageing, (ii) the current role of leisure in daily travel behaviour, and (iii) expected changes in overall mobility patterns.

Observation 1: The ageing society. Population ageing is a global demographic trend. The share of the world population aged 65 years or above is projected to rise from around 10% in 2022 to 16% in 2050 (United Nations, 2022). Switzerland follows the same trajectory and is experiencing a gradual shift towards an older age structure, driven by low fertility and rising life expectancy (Christensen et al., 2009; Kanasi et al., 2016). Population scenarios produced by the Federal Statistical Office (BFS) indicate that the proportion of people aged 65 and over will continue to increase over the coming decades (BFS, 2010, 2025a; Lewis & Ollivaud, 2020). Population ageing is therefore not a temporary fluctuation, but a structural feature of Swiss demography. Figure 1 illustrates this development by comparing the age pyramid in 2025 (lines) with the projected age structure in 2055 (bars).

Observation 2: The scale of leisure. Data from the Swiss Mobility and Transport Microcensus show that leisure has been the most frequent trip purpose in Switzerland in recent years (ARE & BFS, 2021). Pooling the 2015 and 2021 waves, leisure accounts for 37.3% of all daily trips among people aged 6–64 and for 53.5% among older adults aged 65+ (see Figure 2). The second most frequent purpose differs across age groups: among people aged 6–64, work-related travel accounts for 22.3% of daily trips, whereas among people aged 65+ shopping accounts for 26.7%.

Observation 3: An expected increase in leisure. International evidence suggests a structural shift in trip purposes, with non-work and leisure travel gaining relative importance in post-COVID mobility, while shopping increasingly moves online (Wang & De Vos, 2025). This pattern is also consistent with projections for Switzerland. National transport outlooks produced by the Swiss Federal Office for Spatial Development (ARE) point in the same direction. In the Transport Outlook 2050 (ARE, 2021, 2022), the Basis scenario assumes widespread teleworking, the digitalisation of services, and changes in retail. Under these assumptions, leisure trips become more frequent, while trips for other purposes are expected to decline compared with today. The number of leisure trips per

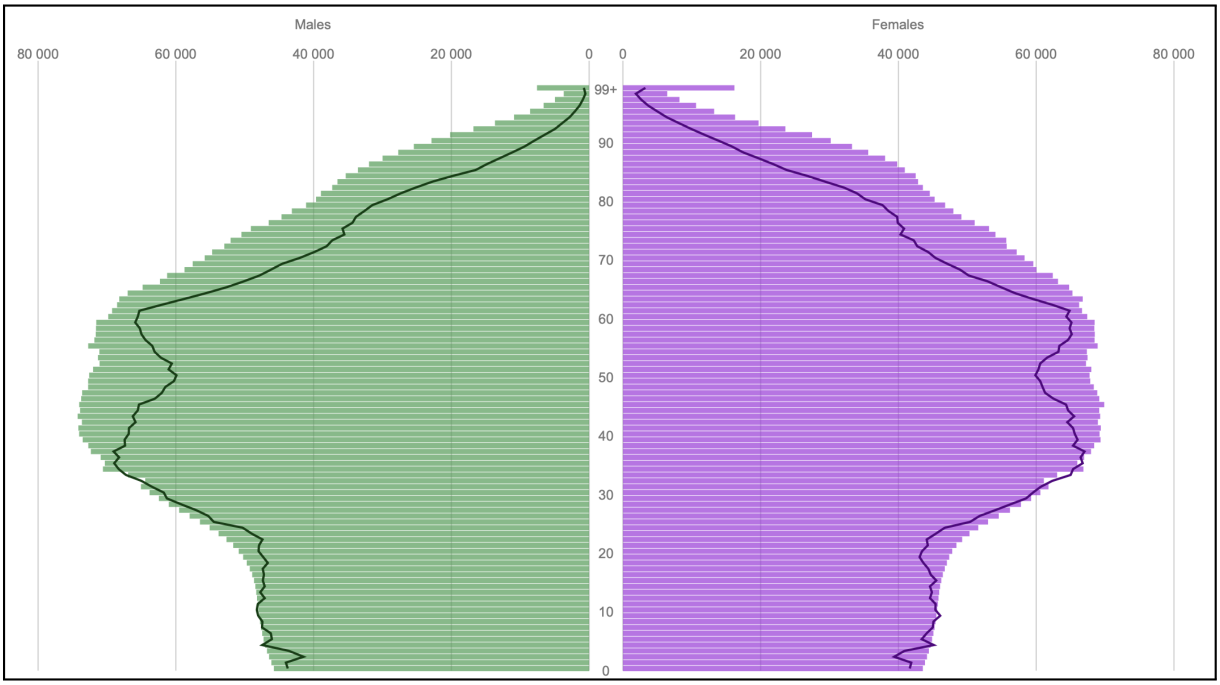


Figure 1: Swiss population pyramid by age and sex in 2025 (lines) and in the 2055 reference scenario (bars). (BFS, 2010, 2025a).

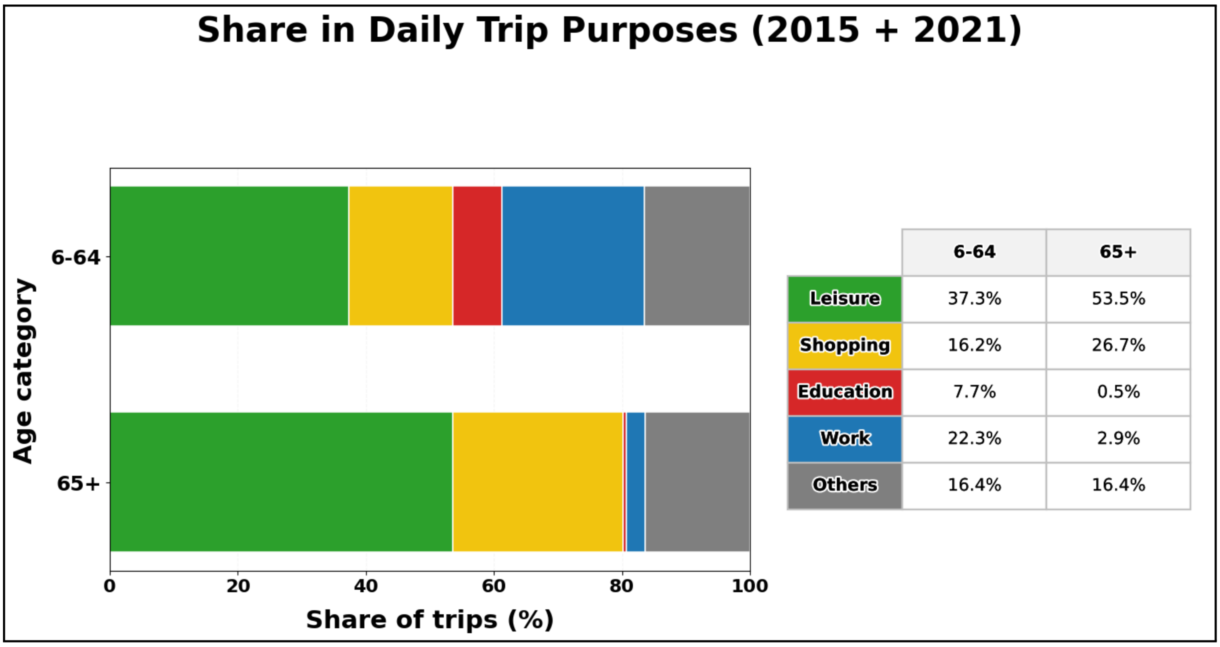


Figure 2: Daily trip purpose shares by age group in Switzerland, pooled Microcensus 2015 and 2021. The category "Others" mainly comprises services/errands and escort trips. (ARE & BFS, 2021).

person and working day is projected to increase by around 15.6% (Table 1). This shift does not necessarily imply that people will travel more in absolute terms; rather, it suggests that a larger share of everyday mobility will be devoted to leisure-related activities.

Taken together, these observations motivate a closer examination of leisure mobility.

Table 1: Relative change in trips per person and working day by purpose in 2050 compared with today, in the Basis scenario of ARE (2022).

Trip purpose	Change [%]
Leisure	**+15.6**
Work	-28.1
Shopping	-5.5
Education	-7.1
Commercial	-20.8

If Switzerland's population is ageing, if leisure already accounts for a particularly large share of travel among older adults, and if leisure is expected to gain further importance in overall mobility patterns, then leisure mobility must be represented more accurately in mobility analysis and travel demand modelling.

1.2 Leisure mobility as a modelling challenge

Leisure refers to activities and trips undertaken during free time, outside mandatory or externally imposed occupations such as work or education (ARE & UNIL, 2018). ARE (2020a) and Ohnmacht et al. (2009) emphasise that, among all trip purposes, leisure is the most difficult to analyse and model.

Leisure is particularly challenging to represent for several reasons. First, it encompasses highly heterogeneous activities with very different distances, durations, and mode choices (ARE & UNIL, 2018; Simma et al., 2001; Spinney & Millward, 2013). A meal at a nearby restaurant, a daily dog walk, a weekly sports activity, and a day trip to the mountains may all be labelled "leisure". Second, the conceptual boundaries of leisure are often blurred and in practice overlap with related notions such as *recreation*, *tourism*, and *bleisure* (Caicedo-Barreth et al., 2020; Mokras-Grabowska, 2018). Third, leisure behaviour is not homogeneous across the population, but can vary across socio-demographic groups and travel contexts. Existing research links age, socio-economic position, employment status, and trip context to differences in leisure mobility behaviour (Buehler et al., 2024; Mahdi et al., 2022; Reece, 2004; Scotti et al., 2024; Spinney & Millward, 2013). Fourth, destination-related behaviour is also shaped by less tangible attributes such as place attachment, destination loyalty, and prior experiences (Li et al., 2023; Moreira & Iao, 2014; Yoon & Uysal, 2005). Finally, many leisure activities are socially coordinated, with decisions made within couples, families, or peer groups rather than by isolated individuals (Arentze, 2015). Taken together, these features make leisure destination choice difficult to represent in a way that is both behaviourally meaningful and operationally tractable.

1.3 Problem statement

From a modelling perspective, leisure is typically represented through destination-choice mechanisms. In many travel demand models, the probability that an individual chooses a destination is formulated within a random-utility framework, in which destinations with higher expected utility are more likely to be selected. In zone-based settings, destination choice utility is commonly structured around two building blocks: (i) an *accessibility* (or impedance) component capturing the effort required to travel from an origin zone i to a destination zone j, and (ii) a destination *attractivity* (or size) term that summarises the amount and appeal of opportunities in zone j. While accessibility varies across origin–destination pairs, attractivity is usually defined as a destination-specific quantity. Intuitively, a destination that is both easy to reach and highly attractive is more likely to be chosen.

A recurring practical limitation in destination choice modelling concerns the specification of the destination-side attractivity term A_j. In many applied models, this term is often represented through relatively simple zonal size terms or aggregate proxies, such as employment, households, or resident population (Clifton et al., 2016; Molloy & Moeckel, 2017). Such representations are operationally convenient, but they provide only a coarse approximation of what makes destinations attractive, especially for discretionary and leisure travel. It therefore becomes difficult to represent the diversity and composition of leisure opportunities across places, and equally difficult to identify and interpret differences in destination-side preferences across population groups or travel contexts. These limitations also arise in operational modelling practice. In the current implementation of SIMBA MOBi, leisure attractivity A_j is composed of a small number of zonal totals. More specifically, it is based on resident population (Pop_j), resident university students ($Stud_j$), and a proxy for leisure-related visitors ($Vis_{\text{leisure},j}$). This visitor proxy is derived from employment in selected leisure-related sectors (see Section 2.3).

Recent work has shown that richer destination-side specifications, including multidimensional indicators and open-data-based measures, can improve the representation of destination attractivity in applied modelling (Klinkhardt et al., 2021; Molloy & Moeckel, 2017; Notelaers et al., 2024). However, these studies have mostly addressed general-purpose travel demand modelling, rather than an operational leisure-specific formulation designed to distinguish population groups and travel contexts. Other leisure-focused destination-choice studies have explicitly shown that leisure is not a homogeneous activity domain, but they often do so by concentrating on specific subsets of leisure. For example, Simma et al. (2001) estimate separate destination-choice models for selected activities such as skiing, climbing and hiking, walking, and swimming. Gramsch-Calvo and Axhausen (2025) focus on restaurants, cafés, bars, and nightclubs to examine how homophilic preferences related to age, income, and cultural background influence leisure venue choice. At the broader end of the literature, tourism and urban-recreation studies describe destination attractivity through a wide range of natural, cultural, recreational, infrastructural, and service-related attributes, or through comprehensive synthetic indices (Biernacka et al., 2020; Krešić & Prebežac, 2011; Singh & Tiwari, 2016; Ziernicka-Wojtaszek & Malec, 2022). These approaches are valuable for identifying relevant dimensions of attractivity, but their breadth can make them too complex and difficult to translate directly into an estimable and operational travel demand model. The resulting challenge is therefore to construct a leisure attractivity formulation that is rich enough to represent leisure opportunities more comprehensively and to support comparisons across population groups and travel contexts, while still remaining sufficiently compact, interpretable, and tractable for operational use in applied travel demand modelling.

1.4 Research objectives and questions

This thesis focuses on daily *domestic leisure*, that is, leisure undertaken within a single day by the resident population within Switzerland, Liechtenstein, and the enclaves of Campione d'Italia and Büsingen am Hochrhein. Trips involving overnight stays are therefore excluded. In addition, *shopping* is treated as a distinct discretionary purpose and is therefore not considered in this thesis.

The overarching aim of the thesis is to develop a richer zone-level leisure attractivity term for destination-choice components in zone-based travel demand models. Operationally, this term is formulated as a leisure attractivity index. The proposed index is intended to provide a more comprehensive representation of territorial leisure opportunities while remaining interpretable and tractable for applied modelling. It is also intended to support the identification and interpretation of heterogeneous destination-side preferences across relevant segments. The approach is developed and evaluated in the Swiss context using observed leisure trips from the Mobility and Transport Microcensus (travel diary survey), with SIMBA MOBi serving as an applied case study (ARE & BFS, 2021, 2023; Scherr et al., 2019).

This thesis does not aim to capture the full complexity of leisure behavioural heterogeneity. Instead, as a pragmatic first step to test whether the proposed formulation can reveal behavioural differences, the analysis focuses on two ex-ante dimensions. The first is *age*, using a threshold at 65+ years to distinguish retirement-age individuals from the rest of the population. This split is primarily motivated by the central role of population ageing in the thesis and by empirical evidence showing that leisure already accounts for a particularly large share of travel among older adults. Accordingly, the 6–64 segment should be interpreted as a broad comparison group rather than as a behaviourally homogeneous category. The second dimension is *travel context*, distinguishing shorter leisure embedded in everyday routines from longer leisure tours that occupy a substantial share of the day. This distinction follows the structure of the Swiss Mobility and Transport Microcensus, in which longer day trips (*Tagesreisen*) are defined as tours with a minimum duration of three hours (see Section 4.2). This dimension is relevant because demographic ageing and retirement are expected to increase the importance of leisure patterns that are less constrained by work schedules. Moreover, longer leisure tours were identified by SIMBA MOBi experts as a particularly relevant blind spot in the current operational representation.

Combining these two dimensions yields four segments s: *Young–Short*, *Old–Short*, *Young–Long*, and *Old–Long*. For each segment, revealed preferences are derived from Microcensus trip records and used to estimate a separate zonal destination-choice model over NPVM traffic zones (Section 4.1) within a multinomial logit framework (Section 5.3). Destination utilities combine an accessibility control term (see Section 4.4) with the proposed zone-level attractivity component.

Concretely, the thesis pursues three interrelated goals:

- **G1 – Constructing a richer attractivity formulation.** Develop a composite leisure attractivity index $A_j(s)$ for traffic zones that captures the territorial context of leisure destinations more comprehensively than simple proxies. The index is based on a limited set of interpretable and tractable supply-side indicators and spatial characteristics, so that it remains of manageable dimensionality for estimation and implementation. Wherever feasible, indicators are derived from open and transferable sources (e.g. OpenStreetMap), complemented by official data where required. The selection of components is grounded in the literature (Chapter 3), and their operationalisation is detailed in Section 4.3 and Section 5.1.

- **G2 – Revealing behavioural differences across segments.** Estimate segment-specific attractivity formulations for the four behavioural segments. Assess whether a richer and multidimensional attractivity specification can reveal differences in destination-side preferences across age groups and travel contexts. Identify which destination-side features are associated with leisure destination choice in each segment.

- **G3 – Evaluating explanatory and operational value.** Evaluate whether the refined attractivity representation improves the explanatory and predictive performance of leisure destination-choice models relative to simpler baseline specifications, including the current operational formulation. This evaluation is carried out both econometrically, through model fit and predictive performance against observed Microcensus data, and operationally, through integration into SIMBA MOBi. In this implementation step, the remainder of the modelling system is kept unchanged.

These goals are translated into the following research questions, which mirror G1–G3:

- **RQ1 – Definition and operationalisation (G1).** How can leisure attractivity $A_j(s)$ be defined and operationalised at the traffic-zone level so as to capture the territorial context and diversity of leisure opportunities?
- **RQ2 – Segment-specific patterns (G2).** To what extent can the proposed leisure attractivity index reveal differences in destination-side preferences across age groups and travel contexts? Which destination-side features emerge as relevant for leisure destination choice in each segment?
- **RQ3 – Added value relative to the baseline (G3).** To what extent does a richer and segment-specific representation of leisure attractivity improve the explanation and prediction of leisure destination choice relative to simpler baseline specifications?

In doing so, the thesis makes three main contributions. First, it develops a structured, flexible, interpretable, and transferable zone-level formulation of leisure attractivity that captures territorial heterogeneity more comprehensively than conventional aggregate proxies while remaining suitable for operational use. Second, it provides empirical evidence on whether and how such a formulation helps reveal and interpret differences in destination-side preferences across behavioural segments. Third, it proposes a transparent and modular workflow to derive indicators, largely from open data, estimate the models, and benchmark the refined specification against baseline alternatives using quantitative metrics. From a practical perspective, the main output is an implementation-oriented attractivity layer that can be integrated into operational modelling frameworks to support a more differentiated representation of leisure destination choice.

The thesis is organised as follows. Chapter 2 introduces the general principles of travel demand modelling and describes in more detail the Swiss modelling framework and the current representation of leisure. Chapter 3 reviews the literature on leisure destination attractivity and competitiveness and derives the conceptual foundations for the attractivity index. Chapter 4 and Chapter 5 present, respectively, the data, indicators, and methodological steps used to construct and estimate the indices. Chapter 6 reports the main empirical findings and the outcomes of integrating the refined $A_j(s)$ into SIMBA MOBi. Chapter 7 discusses the main findings and limitations of the study, outlines directions for future research, and considers implications for practice and policy. Finally, Chapter 8 concludes the thesis.

2 Background: travel demand models

This chapter provides the conceptual background for the rest of the thesis. First, it briefly reviews the core ideas from travel demand modelling that are needed to understand how destination choice and zone-level attractivity are represented (Section 2.1). Second, it presents the Swiss modelling ecosystem, in particular the SIMBA family and the National Passenger Transport Model (Section 2.2). Third, it presents the current leisure attractivity definition used in SIMBA MOBi as the case-study reference for this thesis (Section 2.3).

2.1 Model families and destination choice

Travel demand models are analytical and simulation frameworks used to represent how people generate trips and activity patterns, choose destinations and modes, and collectively load the transport system in order to predict travel demand and passenger flows. They are widely used to assess transport policies, infrastructure projects, and service concepts (Ben-Akiva, 1973). Travel demand models have evolved considerably over the past decades, but this evolution is better understood as a diversification along a few key dimensions than as a replacement of one "generation" by another. Four dimensions are particularly relevant: the level of aggregation (from zonal flows to individual agents), the degree of behavioural realism (from empirical relationships to random-utility and activity-based frameworks), the representation of temporal dynamics (from static average days to microscopic time-dependent simulation), and the use of data (from household surveys to synthetic populations and Big Data) (Anda et al., 2017; Bastarianto et al., 2023; Goulias, 2018; Khademi & Timmermans, 2011). Across these dimensions, different demand model families coexist, complement each other and are often combined in practice. The main families are trip-based four-step models, activity-based models and agent-based models.

Trip-based four-step models and spatial interaction

The traditional point of departure is the trip-based four-step model developed in the 1950s and 1960s (Michigan State Highway Department, 1955; Mitchell & Rapkin, 1954; State of Illinois, 1959). The unit of analysis is the flow between traffic analysis zones (TAZs) (Barceló, 2010; Ghodmare & Yadav, 2021; Ihrig et al., 2024). In its canonical form, the four-step framework comprises (i) *trip generation* (how many trips are produced and attracted by each zone), (ii) *trip distribution* (how many trips go specifically from each origin zone to each destination zone), (iii) *mode choice* (allocating trips to car, public transport, etc.), and (iv) *assignment* (loading flows onto the transport network). (Anda et al., 2017; Bhat & Koppelman, 1999; Goulias, 2018).

In this paradigm, the origin–destination pattern is commonly represented through spatial interaction (gravity-type) models within the trip-distribution step. A classical formulation assumes that the flow T_{ij} between an origin zone i and a destination zone j is proportional to zonal *productions* and *attractions* and decreases with travel impedance:

$$T_{ij} \ \propto\ O_i\, A_j\, f(c_{ij}), \tag{1}$$

where O_i denotes productions at origin i, A_j is an attraction/size term for destination zone j (e.g. population, employment, retail floor space) and $f(c_{ij})$ is a decreasing function of generalised cost c_{ij} (Ghasemi et al., 2021; Kerkman et al., 2017; Wilson, 1971). While this is an aggregate flow formulation, it already mirrors two core building blocks that reappear in disaggregate destination-choice models: an impedance component and a destination-side size or attractivity component.

Despite well-documented behavioural limitations, most notably the treatment of trips as independent events and the limited representation of temporal or household constraints, four-step models remain widely used for strategic planning at national and supranational scales (Ghodmare & Yadav, 2021; Ihrig et al., 2024).

Activity-based models

Beginning in the 1970s and 1980s, several authors argued that trip-based models are misaligned with the way people organise their daily lives. Foundational work in time geography and activity theory emphasised that travel is a derived demand: people move in order to participate in activities distributed over space and time, subject to constraints on time budgets and resources (Axhausen & Gärling, 1992; Hägerstrand, 1970; Kitamura, 1988; Recker et al., 1986). In this view, the appropriate unit of analysis is not the isolated trip but a daily (or weekly) pattern of activities and tours.

Activity-based models implement this shift by aiming to predict, for each individual, which activities are undertaken, in what sequence, where, when, for how long and with which mode (Bhat & Koppelman, 1999; Goulias, 2018). Outputs are complete daily activity–travel schedules that can then be aggregated into flows. Conceptually, activity-based frameworks combine (i) activity generation and participation models, (ii) scheduling and sequencing models for activities and tours, (iii) destination and mode choice models for tours and trips, and (iv) duration models for activities and travel episodes (Bhat & Koppelman, 1999; Chu et al., 2012).

A key behavioural foundation frequently used to operationalise several of these decisions, including destination choice, is discrete choice theory based on random utility maximisation. Following early work by Ben-Akiva and Lerman (1985), Domencich and McFadden (1975), and McFadden (1974) and later syntheses such as Train (2009), the basic premise is that a decision-maker n associates a latent utility U_{ni} with each alternative i in a choice set, and alternatives with higher utility are chosen with higher probability.

Utility is decomposed into a systematic component V_{ni} and a random term ε_{ni}; under standard assumptions on ε_{ni} this yields logit-type models:

$$U_{ni} = V_{ni} + \varepsilon_{ni}. \tag{2}$$

Random-utility models have become a standard behavioural foundation of many modern travel demand models (Ben-Akiva & Lerman, 1985; Train, 2009). In destination-choice applications, utility typically combines a travel-impedance term with a destination-side size or attractivity term (Horni et al., 2008; Jonnalagadda et al., 2001; Scherr et al., 2020). In this sense, disaggregate random-utility models are closely related to the logic of gravity-type spatial interaction models.

Recent work has compared discrete choice models with machine-learning (ML) classifiers for predicting mode or destination choices (Wang et al., 2024b). ML models often outperform classical discrete choice models in predictive accuracy, although the latter retain a strong interpretability advantage (Rahnasto & Hollestelle, 2024). At the same time, systematic reviews highlight that validation, in the sense of assessing predictive performance on holdout or external datasets, is still relatively rare in the discrete choice literature (Parady et al., 2021).

Activity-based models generally offer greater behavioural realism than trip-based frameworks, as they can explicitly represent interdependencies within daily activity–travel patterns. Empirical comparisons between four-step and activity-based models confirm that activity–travel patterns can more accurately reproduce complex trip/tour structures (Zhong et al., 2015). At the same time, activity-based frameworks are more demanding in terms of data, computation and calibration (Haghighi & Miller, 2025).

Agent-based transport simulation and microsimulation

Agent-based transport models (also referred to as multi-agent or microscopic simulations) emerged from the late 1990s and 2000s as an extension of activity-based modelling in which both demand and supply are represented in a fully dynamic, individual-level framework (Balmer et al., 2004; Horni et al., 2016; Nagel et al., 1997; Raney et al., 2003). Building on activity-based concepts, these models represent each traveller as an autonomous agent with a daily activity–travel plan and simulate the execution of these plans on a detailed transport network over time (Anda et al., 2017; Bastarianto et al., 2023; Goulias, 2018). Because network loading is dynamic, congestion, crowding and interactions between agents can emerge endogenously.

Frameworks such as TRANSIMS and MATSim follow a common pipeline. First, a synthetic population is generated. Second, activity schedules and initial mode and destination choices are created, often using activity-based components. Agents are then simulated on the transport network (Horni et al., 2016; Nagel et al., 1997). Many implementations include iterative replanning or learning mechanisms in which agents adjust their behaviour in response to experienced conditions. Such models can be applied at both metropolitan and national scales, and they have also been used in more targeted applications that couple activity-based demand with microscopic simulation for specific policy or service questions (Armellini et al., 2021; Ziemke et al., 2019).

Agent-based transport models are particularly attractive for analysing policies that depend on dynamic network conditions, interactions between travellers and detailed service design (e.g. congestion pricing, new mobility services, crowding effects). However, they are computationally intensive and require consistent integration of modules for synthetic population generation, activity-based demand, route choice, learning and network simulation (Anda et al., 2017).

Zonal attractivity and data foundations

Across all model families, the representation of destination choice hinges on how the spatial distribution of opportunities is captured. In aggregate four-step models, this is done through zonal productions and attractions. In disaggregate activity- and agent-based frameworks, a set of indicators are used as explanatory variables in destination-choice utilities. In all cases, the need to translate heterogeneous spatial information into tractable measures of zonal attractivity remains central. The specification and estimation of such components depend on the available data used to observe flows or revealed destination choices. Classic household travel surveys and censuses remain indispensable because they provide information on trip purposes, destinations chosen, socio-demographics and full daily patterns (Chu et al., 2012; Goulias, 2018). At the same time, passive data sources such as public transport smart cards and mobile phone records offer greater spatial and temporal coverage (Anda et al., 2017; Jeewanthi & Kumarage, 2025).

2.2 Swiss national travel modelling ecosystem

Switzerland has a diversified travel modelling landscape involving national public authorities, transport operators and regional administrations. Some of its main frameworks partly rely on shared spatial references, data foundations and modelling inputs. At the federal level, the Federal Office for Spatial Development (ARE) maintains the National Passenger Transport Model (NPVM), which serves as the national reference framework for strategic passenger transport planning (ARE, 2020a, 2020b, 2025, 2026). Within the Swiss Federal Railways (SBB), the SIMBA family provides an environment for demand forecasting and service evaluation, combining the macroscopic rail model SIMBA Bahn with the multimodal microscopic model SIMBA MOBi (Scherr et al., 2020; Scherr et al., 2018, 2019). At the regional level, cantons operate their own multimodal models, such as the *Gesamtverkehrsmodell* (GVM-ZH) for the Canton of Zurich (Amt für Mobilität, Kanton Zürich, 2025). The following sections briefly introduce NPVM and SIMBA from a conceptual perspective focusing in particular on destination choice and zonal attractivity.

National Passenger Transport Model (NPVM)

The NPVM is Switzerland's national multimodal passenger transport model and serves as the federal reference framework for strategic transport planning (ARE, 2020b, 2025, 2026). In terms of the model families introduced previously, it broadly follows the logic of a trip-based four-step model operating on zonal flows, while incorporating a relatively rich behavioural structure. Demand is segmented by trip purpose and by behaviourally similar population groups, while destination and mode choices are modelled jointly using discrete-choice methods. Spatially, NPVM operates on a system of 7,978 traffic zones covering Switzerland and 710 adjacent foreign zones. These traffic zones constitute the basic spatial unit used in this thesis (see Section 4.1).

For the purposes of this thesis, the key point is that NPVM combines behavioural differentiation on the demand side with a more aggregate representation of destination attractivity. For each origin–destination purpose relation (for example home–work or home–leisure), the model distinguishes behaviourally similar population groups based on characteristics such as age, employment status, and mobility tools (ARE, 2020a). In line with the gravity-type logic introduced in Equation (1), zonal flows between origin i and destination j depend on origin-side productions, destination-side attractivity, and travel impedance. In the case of leisure, destination attractivity is represented through common zone-level attraction terms that vary by leisure distance, distinguishing short-distance leisure trips ($<$ 10 km) from long-distance leisure trips ($>$ 10 km) (ARE, 2020a). This distance-based differentiation is relevant for the present thesis, which builds on the same intuition while adopting a different segmentation logic based on age group and tour duration.

Swiss Federal Railways: the SIMBA framework

Within SBB's modelling landscape, *SIMBA* (***S**tandardisierte, **I**ntegrierte **M**odellierung und **B**ewertung von **A**ngebotskonzepten*, i.e. standardised, integrated modelling and evaluation of service concepts) denotes the corporate modelling framework used for demand forecasting, planning, and service evaluation. It consists of two complementary components. The first, *SIMBA Bahn* (*Bahn* meaning rail), is a macroscopic model focused on station-to-station rail demand. The second, *SIMBA MOBi* (***Mo**bilität, **B**ahn und **i**ntermodal*, i.e. mobility, rail, and intermodal transport), is a multimodal microscopic model that simulates the daily mobility of individual residents across transport modes and trip purposes. The two models are used in combination within SBB's planning environment. In terms of the model families introduced in Section 2.1, SIMBA Bahn broadly follows the logic of trip-based macroscopic demand and assignment models, whereas SIMBA MOBi combines activity-based demand generation with agent-based network simulation.

SIMBA MOBi covers the whole of Switzerland, including Liechtenstein and the enclaves of Campione d'Italia and Büsingen am Hochrhein, and also represents cross-border flows from neighbouring regions whose trips enter the Swiss network. It simulates a full average weekday for a synthetic population of individual agents. Although SIMBA MOBi is microscopic, some model components are still defined at an aggregate zonal level. For this reason, the model inherits the NPVM traffic-zone system as a common spatial reference. On the demand side, daily activity–travel plans are generated through econometric and discrete-choice models. These decisions generally depend on agents' socio-demographic characteristics, mobility resources, and activity constraints. On the supply side, SIMBA MOBi represents the main elements of the Swiss transport system that agents can use during the simulated day. These include a nationwide road network, a public transport network based on the full timetable, and simplified walking and cycling representations. In addition, the model uses multimodal origin–destination matrices between NPVM traffic zones summarising travel times, distances, transfers, and service frequencies. The execution of daily plans on the multimodal network is carried out in MATSim, an agent-based transport simulation platform (Horni et al., 2016).

Internally, SIMBA MOBi is organised into three main modules: *MOBi.SynPop*, which provides the synthetic population and spatial reference system; *MOBi.Plans*, which generates daily activity–travel schedules and behavioural choices; and *MOBi.Sim*, which executes these plans in a dynamic agent-based simulation. The output of one module constitutes the input for the next.

MOBi.SynPop provides the synthetic population and spatial reference system used by SIMBA MOBi. It builds on the national *SynPop* (ARE & SBB, 2024), jointly developed by ARE and SBB from official registers. In particular, STATPOP provides information on residents and households (BFS, 2024), while STATENT provides information on enterprises and jobs (BFS, 2023). SynPop represents these elements across Switzerland, Liechtenstein, and the enclaves. In this sense, the synthetic population is designed to be statistically representative of the resident population, while establishments and activity locations provide a spatially explicit basis for potential destinations. Each synthetic person is also associated with a home location that serves as the spatial anchor of daily activity–travel schedules. In SIMBA MOBi, these data are used at both the individual and the zonal level. At the microscopic level, synthetic persons carry attributes such as age, household membership, employment and education status, car availability, and public transport subscription. Establishments and activity locations provide the spatial basis for potential destinations such as workplaces, schools and leisure facilities. At the zonal level, persons, households, and establishments are aggregated to NPVM traffic zones.

This allows the derivation of structural indicators relevant for travel demand modelling, such as resident population per zone and estimated visitors.

MOBi.Plans implements the activity-based demand component of SIMBA MOBi. For each synthetic person, it generates a complete 24-hour activity–travel schedule consisting of activities and travel legs organised in tours. In doing so, it transforms the socio-demographic and spatial information provided by MOBi.SynPop into daily plans with assigned activities, destinations, modes, and timing. Activities are represented using the same main categories as in NPVM. Primary activities include work (W) and education (E). Secondary activities include business (B), leisure (L), shopping (S), accompany (A), and other (O). Home (H) serves as the anchor activity. Conceptually, MOBi.Plans follows a sequence of behavioural decisions. Longer-term choices, such as mobility-tool ownership and primary work or education locations, provide the starting point. Conditional on these, the model generates daily activity participation, tour structure, timing, destination, and mode choices. Additional rule- and optimisation-based steps ensure temporal and spatial consistency of the final schedules.

These behavioural models follow the random-utility framework introduced in Section 2.1. For the purposes of this thesis, the most relevant aspect is that destination choice for secondary activities is handled within this module. The output of MOBi.Plans is a *plan file* for each agent: a time-ordered sequence of activities and trips with assigned locations, modes, and departure times.

MOBi.Sim is the dynamic simulation engine of SIMBA MOBi. It takes the daily activity–travel plans generated by MOBi.Plans, together with additional exogenous trips from abroad, and executes them on the detailed multimodal network using MATSim and SBB-specific extensions (Horni et al., 2016; Scherr et al., 2019). In this step, agents move through the transport system, experience realised travel times and network conditions, and interact with congestion, capacities, and timetable constraints. Execution, scoring, and replanning are iterated until the system reaches a stable state. In each iteration, agents evaluate the realised quality of their plans, and a subset of them is allowed to adjust network-related choices such as routes. This iterative process captures feedback between individual behaviour and network performance. The simulation produces indicators such as modal split, link and station loads, travel times, and distances.

2.3 Leisure destination choice within SIMBA MOBi

Destination choice for secondary activities determines where non-mandatory activities such as shopping or leisure take place. In SIMBA MOBi, this choice is implemented in two stages: first, a traffic zone is selected; second, a concrete facility is chosen within that zone. The overall structure of the destination-choice model is the same across secondary activities, but the corresponding utility functions differ by activity type through specific attractivity terms. This thesis focuses on the first stage, i.e. zonal leisure destination choice.

Destination choice in SIMBA MOBi follows the random-utility framework introduced in Equation (2). In the present context, utility is defined for each origin–destination pair, since the probability of visiting a destination depends not only on the opportunities available at destination j, but also on how easily it can be reached from origin i. For leisure destination choice, the systematic utility of choosing destination zone j from origin zone i can therefore be written in simplified form as

$$V_{ij} \;=\; \alpha\, \mathrm{Accessibility}_{ij} \;+\; \beta\, \mathit{Attractivity}_j^{\mathrm{leisure}}, \tag{3}$$

The attractivity term is discussed in the remainder of this section, while the accessibility measure, referred to in SIMBA MOBi as expected multimodal utility (EMU), is discussed in Section 4.4. In the operational workflow, the utility can additionally include correction terms (λ_j and λ_{ij}), used to align simulated choices with aggregate targets.

Under the standard logit assumption for the random utility components, this specification yields the destination-choice probability

$$P_{ij} \;=\; \frac{\exp(V_{ij})}{\sum_k \exp(V_{ik})}, \tag{4}$$

which defines a proper probability distribution over all feasible destination zones.

For secondary activities, destination choice is often embedded in trip chains that also involve a primary location such as work or school. Leisure destination choice is therefore not conditioned only on the home location, but also on the primary location of the tour. In SIMBA MOBi, this is handled through a *rubber banding* mechanism, which blends home-based and primary-location-based destination probabilities:

$$P_j^{\mathrm{RB}} \;=\; \gamma\, P_{hj} \;+\; (1-\gamma)\, P_{pj}, \qquad \gamma \in [0,1], \tag{5}$$

where P_{hj} denotes the probability of choosing destination j when conditioning on the home location h, and P_{pj} denotes the corresponding probability when conditioning on the primary location p.

In the current SIMBA MOBi specification, the zonal attractivity term for leisure is defined as a weighted combination of three zonal variables:

$$A_j^{\mathrm{leisure}} \;=\; \beta_1\, \mathrm{Pop}_j \;+\; \beta_2\, \mathrm{Vis}_{\mathrm{leisure},j} \;+\; \beta_3\, \mathrm{Stud}_j, \tag{6}$$

where Pop_j is the resident population in zone j, $\mathrm{Vis}_{\mathrm{leisure},j}$ is a zonal indicator of leisure-related visitors, and Stud_j is the number of resident university students. In the operational specification, the corresponding coefficients are $\beta_1 = 0.2382$, $\beta_2 = 2.3820$, and $\beta_3 = 0.2000$.

A central element is the leisure visitor variable $\mathrm{Vis}_{\mathrm{leisure},j}$, which is constructed from establishment-level information and then aggregated to NPVM traffic zones. The starting point is STATENT (BFS, 2023), which provides information on each enterprise u, including its location, NOGA category, i.e. the Swiss General Classification of Economic Activities (BFS, 2008), and the number of full-time equivalents (FTE). Based on the NPVM methodology, each leisure-related activity category is assigned a conversion factor that translates employment into an approximate number of visitors (ARE, 2020a). In simplified form, the number of visitors associated with activity type A and enterprise u can be written as:

$$Vis_{A,u} \;=\; \mathrm{FTE}_u \cdot b_A, \tag{7}$$

where b_A denotes the activity-specific conversion factor (visitors per FTE per year). Activities that are assumed to attract more visitors are assigned higher visitor potentials. Table 2 shows a small selection of such conversion factors.

Table 2: Example conversion factors b_A for leisure-related NOGA sectors.

Activity type	NOGA example	b_A (visitors/FTE/year)
Cinemas	5914	10 980
Gastronomy (restaurants, bars)	5610	3 500

At the zonal level, leisure visitors are obtained by summing these estimated visitor volumes over all relevant facilities located in zone j. For illustration, a cinema with 10 full-time equivalents and a conversion factor of 10 980 generates an estimated 109 800 visitors per year, or around 300 per day. Once a destination zone has been selected, the model further assigns the choice to a specific facility within that zone. Facilities with higher visitor volumes receive a higher probability of being selected.

This thesis proposes a new attractivity function as an alternative to the specification in Equation (6). While the proposed reformulation is developed and tested within the SIMBA MOBi case study, its underlying logic is not specific to this framework and is, in principle, transferable to other discrete-choice settings. SIMBA MOBi is therefore used here as an operational example of how zonal attractivity enters destination-choice modelling. For comparability with the existing implementation, this thesis retains the same accessibility logic used in SIMBA MOBi in the generic utility specification, as well as the subsequent facility-level discretisation step.

3 Review: leisure destination attractivity

This chapter reviews the literature on destination attractivity with the specific aim of identifying which destination-side features may contribute to a zone's leisure attractivity. These include, for example, facilities, services, amenities, environmental qualities, and derived indicators. In this sense, the purpose of the chapter is to extract from the literature a conceptually grounded set of observable leisure features that can later be quantified and operationalised in the specification of the attractivity index $A_j(s)$ (Section 4.3).

The chapter draws on evidence from a broad range of modelling traditions. These include discrete choice models of leisure destination choice (Morey et al., 1991; Pozsgay & Bhat, 2001; Simma et al., 2001); gravity-type and spatial-interaction approaches (Marrocu & Paci, 2013; Massidda & Etzo, 2012); composite indices of recreational potential or destination competitiveness (Paracchini et al., 2014; Schirpke et al., 2018); and data-driven studies exploiting mobile-phone or GPS traces and machine-learning techniques (Chen et al., 2024; Zhang et al., 2019). Although these approaches differ, they often point to similar underlying destination-side features. In this thesis, these features are later combined into a zonal leisure attractivity specification. In simplified form, the logic is that attractivity can be expressed as:

$$A_j(s) = \sum_{h=1}^{H} \beta_h(s)\, F_h(j), \tag{8}$$

where $F_h(j)$ denotes the leisure-related features of zone j, and $\beta_h(s)$ captures their segment-specific relevance for behavioural segment s. In the empirical part of the thesis, the features are operationalised through spatial indicators, while the corresponding coefficients are estimated from revealed destination choices within a discrete choice framework.

The empirical focus of this thesis is on leisure trips undertaken within the daily activity space and without overnight stays. At the same time, this review also draws, with caution, on the broader but closely related literatures on recreation and tourism whenever they help identify destination-side features. This broader perspective is justified because leisure, recreation and tourism are conceptually overlapping rather than neatly separable domains (Mokras-Grabowska, 2018; Slak Valek & Fotiadis, 2018). In practice, several activities considered part of everyday leisure are also examined in tourism research, often through similar destination attributes and motivational mechanisms (Slak Valek & Fotiadis, 2018).

A related difficulty, already highlighted in Section 1.2, is the wide heterogeneity of activities that fall under the label of leisure. As a result, leisure activities are difficult to classify consistently into a limited number of meaningful domains. The aim of this thesis is not to resolve this conceptual problem in full, but rather to define a finite set of leisure domains that can guide the operationalisation of spatial features while still covering a category as broad and heterogeneous as leisure. To support this step, the thesis draws on the leisure-purpose categories recorded in the Swiss Mobility and Transport Microcensus (Section 4.2), while also taking inspiration from classification frameworks proposed in the literature, such as the Catalogue of Leisure Activities (CaLA) (Pospíšil et al., 2022).

The chapter is organised as follows. It first introduces key conceptual lenses on destination attractivity (Section 3.1). It then synthesises the empirical literature on dimensions relevant to leisure destination choice (Section 3.2). Finally, it briefly reviews how leisure behaviour is segmented in the literature (Section 3.3).

3.1 Conceptual lenses on destination choice

Destination attractivity is a multidimensional construct that has been approached from different theoretical angles in the tourism and leisure literature. This subsection reviews four complementary conceptual lenses through which destination attractivity can be understood. Taken together, they show how tangible attributes such as infrastructure, facilities, services, and environmental quality interact with less tangible elements such as motivations, lifestyles, symbolic meanings, and emotional bonds in shaping destination choices.

Push–pull frameworks. Push–pull frameworks interpret destination choice as the result of an interaction between internal drivers on the decision-maker side and external attributes on the destination side (Correia et al., 2013; Crompton, 1979; Leong et al., 2015; Yoon & Uysal, 2005). Push factors refer to the preferences, needs, and constraints that lead individuals to seek leisure or tourism experiences, for example escape from routine, relaxation, self-actualisation, or social interaction (Crompton, 1979; Yoon & Uysal, 2005). Pull factors, by contrast, refer to the concrete characteristics of destinations that make them attractive in relation to these preferences, such as landscape and climate, cultural and historical resources, events, safety, facilities, and hospitality (Correia et al., 2013; Yoon & Uysal, 2005). In this view, destinations are chosen because their attributes match the preferences and needs of potential visitors (Yoon & Uysal, 2005). For the present thesis, this framework is useful because it closely mirrors the logic of Equation (8). In simplified terms, the pull side corresponds to the observable leisure features offered by a zone, i.e. $F_h(j)$, whereas the push side helps explain why the same set of features may be valued differently across behavioural segments, as reflected in the segment-specific coefficients $\beta_h(s)$.

Resource-based and competitiveness views. A second lens treats destination attractivity as emerging from the combination of multiple local resources rather than from a single attribute (Crouch, 2011; Dwyer & Kim, 2003). These resources can be understood as including three broad types of elements: endowed resources, such as landscapes, lakes, and local culture; created resources, such as leisure facilities, museums, and sports infrastructure; and supporting resources, such as general infrastructure, service quality, and destination management (Crouch, 2011; Dwyer & Kim, 2003; Heath, 2003). Empirical studies show that both natural and built resources contribute to recreational potential, and that infrastructure and service quality influence how effectively these resources can be used (Świdyńska & Witkowska-Dąbrowska, 2021). For the present thesis, this perspective is useful because it suggests that destination attractivity should be represented as a bundle of complementary supply-side features $F_h(j)$, rather than as a single proxy.

Behavioural and lifestyle perspectives. Behavioural and lifestyle approaches emphasise that leisure mobility is embedded in broader patterns of values, lifestyles, and social relationships. Rather than treating all individuals as homogeneous decision-makers, these perspectives distinguish between mobility styles or lifestyle groups that differ in their preferred activities, travel modes, and spatial practices (Lanzendorf, 2002; Ohnmacht et al., 2009). For the present thesis, this perspective is mainly relevant because it suggests that destination-side features $F_h(j)$ are unlikely to be valued uniformly across individuals, but may instead be interpreted differently across behavioural or lifestyle segments (see also Section 3.3).

Place attachment, image and loyalty. A fourth strand of literature focuses on the emotional and symbolic dimensions of attractivity, conceptualised through notions of place attachment, destination image, and loyalty (Moreira & Iao, 2014; Pan et al., 2021). Research shows that people can develop strong emotional, cognitive, and functional bonds to particular places, often as a result of repeated leisure experiences, social relations, and meaningful events (Arentze, 2015; Reitsamer et al., 2016). These bonds influence future destination choices: tourists and residents attached to a place are more likely to revisit it, recommend it to others, and remain loyal even when competing destinations offer similar or superior attributes (Li et al., 2023; Moreira & Iao, 2014; Yoon & Uysal, 2005). In the modelling framework of this thesis, such dimensions are not introduced as explicit components of $A_j(s)$, mainly because they are difficult to proxy through spatial indicators. They nevertheless provide an important conceptual reminder that $A_j(s)$ is an inherently partial representation of attractivity, and that part of the residual variation in leisure flows may reflect these subjective dimensions.

3.2 Dimensions of leisure attractivity

This section reviews the empirical literature on the destination-side attributes that most consistently shape leisure destination choice. The aim is to identify a finite set of dimensions or domains of leisure attractivity that recur across studies and can plausibly be interpreted as components of zonal leisure attractivity $A_j(s)$. In the remainder of the chapter, these dimensions are referred to as *features* or *feature blocks* and are labelled $F_1, F_2, \ldots$ for convenience. For clarity, the review distinguishes between domain-specific leisure features and broader structural features, both summarised in Table 3. In the empirical implementation in Chapter 4, each feature may be represented by one, several or, in some cases, no direct indicators, depending on data availability and operational feasibility.

Table 3: Feature groups of leisure attractivity.

Domain-specific features	**Overall structural features**
F1: Gastronomy	F7: Diversity
F2: Visits and social activities	F8: Density
F3: Outdoor recreation	F9: Monetary cost
F4: Cultural activities	F10: Supporting infrastructure and services
F5: Sport activities	F11: Spillovers (spatial autocorrelation)
F6: Other leisure domains	

Domain-specific supply features

A first set of dimensions concerns the concrete types of leisure opportunities available in a zone. In this thesis, these are grouped into six domains. Together, they provide a structured way of organising the main forms of leisure supply while remaining sufficiently broad to be operationalised through spatial indicators.

F1: Gastronomy. The first domain comprises restaurants, bars, cafés, and canteens, i.e. places where food and drinks are consumed outside the home. The broader literature similarly treats culinary experiences as a core component of urban attractivity and tourism competitiveness, both in terms of basic provision and as a carrier of local identity and image (Öner & Klaesson, 2017; Ramos & Pinto, 2024; Sirkis et al., 2022). The gastronomy feature $F_1(j)$ can therefore be proxied by indicators that capture both the *quantity* and the *quality or capacity* of gastronomic supply. Typical measures include the number of restaurant, bar, and café POIs, as well as seating capacity where available. Opening hours are also relevant, since they determine how often such places are actually usable within daily schedules (Liu et al., 2021). Online ratings and review volumes may provide an additional proxy for perceived quality and reputation (Ramos & Pinto, 2024).

F2: Visits and social activities. The second domain covers visits to friends and relatives and other informal social gatherings, which are often organised around private homes rather than publicly observable leisure opportunities. Research on leisure mobility highlights the importance of social networks and social influence in shaping leisure travel, and shows that visiting friends and relatives represents a non-negligible component of leisure mobility (Ohnmacht et al., 2009; Schlich et al., 2004). Because these destinations are typically private households or informal meeting places, they are difficult to observe directly in standard POI datasets and other spatial supply data sources.

F3: Outdoor recreation. The third domain covers non-sport outdoor activities such as walking in parks, easy hikes, and the general enjoyment of green and blue spaces. Ecosystem-services and recreation studies consistently show that naturalness, the presence of water, protected areas, and access to forests and trails are key drivers of outdoor recreation potential (Paracchini et al., 2014; Schirpke et al., 2018; van Zanten et al., 2016; Willibald et al., 2019). The availability of parks and green spaces also shapes leisure-related destination choices (Biernacka et al., 2020; Sugiyama et al., 2010). Accordingly, the outdoor feature $F_3(j)$ can be proxied through a combination of green and blue areas (parks, forests, lakes, rivers), and measures of access to walking and hiking infrastructure such as trail length.

F4: Cultural activities. The fourth domain groups cultural and entertainment activities such as museum and gallery visits, theatres, concert halls, and cinemas. Destination competitiveness models and urban cultural policy literature identify cultural resources and events as core attractors that can draw visitors from large catchment areas and contribute to urban regeneration (Crouch, 2011; Krešić & Prebežac, 2011). Empirical studies likewise show that the number of museums and heritage sites, as well as cultural expenditure and event intensity, are positively associated with tourism and leisure (Marrocu & Paci, 2013; Massidda & Etzo, 2012; Sirkis et al., 2022). The feature $F_4(j)$ can therefore be proxied by counts of cultural facilities. Online ratings and review volumes may further serve as proxies for perceived cultural quality and local reputation (Ramos & Pinto, 2024).

F5: Sport activities. The fifth domain captures leisure activities involving active sport, such as fitness, gym workouts, swimming, team sports, and specialised sports facilities (e.g. tennis courts or ice rinks) (Simma et al., 2001; Spinney & Millward, 2013). The active-sport feature $F_5(j)$ can be proxied by the number and capacity of sports facilities (Biernacka et al., 2020; Liu et al., 2021).

F6: Other leisure domains. Finally, a residual domain groups leisure activities that are conceptually distinct but sparsely represented in mobility surveys or POI data when taken individually. Examples include wellness and health-oriented leisure (spas, thermal baths, or wellness centres), religious and spiritual visits (places of worship, cemeteries), amusement parks and funfairs, and youth centres or social clubs (Öner & Klaesson, 2017; Sirkis et al., 2022). Given their heterogeneity, these activities are grouped into a single feature $F_6(j)$, which can be proxied in the same way as the previous domains.

Overall structural features

Beyond the specific leisure categories discussed above, the literature highlights a set of structural features that shape the overall attractivity of places. These features do not correspond to single activity types, but cut across them. In the $A_j(s)$ framework, they are grouped into five components.

F7: Diversity. A first structural dimension is the functional and environmental diversity of places (Liu et al., 2021; Zhang et al., 2019). Similarly, ecosystem-service research finds that landscape diversity, for example the co-presence of forests, agricultural land, water bodies and settlements, plays a role in outdoor recreation potential (Chen et al., 2024; Schirpke et al., 2018). $F_7(j)$ can be measured through diversity of leisure-related POI categories or land-use mix indicators.

F8: Density. A second structural dimension concerns the *density* of opportunities, i.e. how many relevant POIs or facilities are available per unit of area or per capita (Zhang et al., 2019). Urban mobility studies also find that dense areas support more frequent leisure walking and social activities (Liu et al., 2020). $F_8(j)$ can be captured as *POI per area* or *POI per resident* for the relevant leisure categories. This feature distinguishes structurally "rich" zones, with many opportunities within short distances, from sparsely equipped ones.

F9: Monetary cost. A third cross-cutting dimension is the *monetary cost* associated with leisure trips and activities. Destination-choice studies in tourism generally find that higher prices for transport and services reduce flows, all else being equal (Marrocu & Paci, 2013; Morley, 1994). In the present context, monetary cost is therefore considered as a potentially relevant structural feature, although its role may be more difficult to capture for everyday leisure trips than for longer-distance tourism. Where available, $F_9(j)$ can be approximated through local price levels or related cost proxies (Marrocu & Paci, 2013).

F10: Supporting infrastructure and services. A fourth structural feature concerns basic comfort services and organisational features that make leisure settings workable in everyday life. Tourism and urban-attractivity studies show that the availability of services is positively associated with visits (Singh & Tiwari, 2016; Sirkis et al., 2022; Świdyńska & Witkowska-Dąbrowska, 2021). $F_{10}(j)$ can be represented as a composite feature summarising the presence of basic comfort services. This component captures the enabling role of infrastructure: it does not create leisure opportunities on its own, but enhances the usability of existing attractions.

F11: Spillovers (spatial autocorrelation). Finally, numerous studies stress that leisure and tourism attractivity is *not* a purely local property: destinations interact with their neighbours, generating spatial spillovers and autocorrelation effects. Origin–destination models for tourism flows find significant spatial dependence both at origin and destination, implying that demand in one region is influenced by the attractivity of nearby regions (Deng & Athanasopoulos, 2011; Marrocu & Paci, 2013). In the present framework, these mechanisms are not interpreted as an additional feature of $A_j(s)$ itself. Rather, spatial spillovers are treated as a robustness issue and tested separately (see Section 6.3).

3.3 Segmentation of leisure mobility behaviour

The literature discusses differences in leisure mobility behaviour along several dimensions, including age and life-cycle, socio-economic position, gender, health and accessibility constraints, broader lifestyle profiles, and trip context. Taken together, these studies suggest that leisure participation, destination choice, travel distances and activity types are not distributed uniformly across the population.

Age and life-cycle are common segmentation variables in the literature on leisure mobility and tourism, as they are often associated with differences in trip frequency, distance, and activity participation (ARE & UNIL, 2018; Buehler et al., 2024; Reece, 2004). At the same time, chronological age alone does not appear sufficient to explain the full diversity of leisure behaviour (Mathur et al., 1998; Ohnmacht et al., 2009). This is particularly evident within older populations, where senior tourism research points to substantial internal heterogeneity, with segments differing in their motivations, preferences, and activity profiles (Buehler et al., 2024; Mathur et al., 1998; Shoemaker, 1989).

Socio-demographic factors further stratify leisure demand. Income also appears relevant, as socio-economic resources can shape leisure mobility opportunities and constraints (Mahdi et al., 2022; Töger et al., 2023). Gendered differences also emerge in the types of leisure environments and facilities users consider attractive, at least in the context of urban green spaces (Wang et al., 2024a). Accessibility and environmental quality also matter for leisure participation, especially where the usability of nearby open spaces is more important than mere proximity (Sugiyama et al., 2010). Daily schedules and workday conditions can also influence leisure-space use and the types of settings people are more likely to visit (Wang et al., 2024a).

The literature also distinguishes between residents and tourists, and between short, everyday leisure and longer trips. Residents tend to be more strongly associated with dining and shopping activities, whereas tourists are more oriented towards sightseeing and vacation- or health-related destinations (Zhang et al., 2025). Related work on visitor heterogeneity further distinguishes between same-day and overnight travel, suggesting that different trip types respond to different destination attributes and temporal patterns

(Scotti et al., 2024). Studies of sports and recreation similarly show that everyday activities such as walking or exercising are generally undertaken at shorter distances, whereas more specialised activities such as golf or skiing involve broader travel sheds (Spinney & Millward, 2013).

4 Data and operationalisation

This chapter explains how the empirical ingredients required for the destination-choice models are operationalised. The aim is to represent leisure destination choice through an estimated utility function defined over alternative zonal destinations. A destination-choice model requires three main elements: (i) a set of feasible alternatives, (ii) observations of realised choices, and (iii) a set of attributes entering the utility function, describing both the alternatives themselves and the origin–destination relation. As the set of feasible destinations, this thesis uses the NPVM traffic zones developed for the Swiss national modelling framework (Section 4.1). The empirical model therefore represents destination choice between an origin traffic zone i and a destination traffic zone j, where origin and destination may also coincide. Observed choices are derived from the Swiss Mobility and Transport Microcensus (MTMC), from which domestic leisure trips with geocoded origins and destinations are extracted. The representation of destinations is based on recurring destination-side features identified in the literature and operationalised as zone-level supply indicators (Section 4.3). A more parsimonious subset is later selected for model estimation (Section 5.1). In addition, origin–destination accessibility is represented following the general accessibility logic of SIMBA MOBi (Section 4.4). All spatial datasets are harmonised within a GeoPackage-based workflow and projected to the Swiss reference coordinate system CH1903+ / LV95 (EPSG:2056).

4.1 Unit of analysis: NPVM traffic zones

This thesis adopts NPVM traffic zones as the unit of analysis. In principle, the same framework could be applied to other spatial units (e.g. communes, grid cells, or tailored catchment areas), as long as observed trips and spatial indicators can be consistently aggregated to that level.

NPVM traffic zones are functional traffic analysis zones rather than purely administrative units. They are designed to represent coherent mobility units with broadly comparable magnitudes of *structural elements* (in practice, residents plus full-time equivalents), while respecting major physical and network barriers (ARE, 2017). The current system refines the earlier zonal structure where needed in order to keep the number of structural elements per zone within a relatively narrow range. Boundaries are preferentially aligned with major transport infrastructure and natural separators (e.g. railways, primary roads, rivers and lakes). This process produces finer zones in dense urban and agglomeration areas and fewer, larger zones in rural and alpine regions (ARE, 2017).

The choice of NPVM traffic zones offers two main advantages for this thesis. First, it is fully consistent with the Swiss national modelling framework (see Section 2.2). Second, NPVM zones are designed to be broadly comparable in terms of structural size. This helps reduce distortions that would arise if very small and very large spatial units were compared directly. As a result, differences in destination utility can be interpreted more plausibly as reflecting variation in leisure opportunities rather than artefacts of zoning.

The analysis uses the *Verkehrszonen 2017* dataset (VM-UVEK, 2017). The model space retained in this thesis comprises 7,978 zones and includes Switzerland, Liechtenstein, and the two foreign enclaves of Büsingen am Hochrhein and Campione d'Italia, which are retained at municipality level. These special cases, together with traffic zones along large lakes that can be artificially extensive in surface area, occasionally require specific treatment in the construction of spatial indicators. While NPVM also defines 710 zones in neighbouring countries for cross-border flows, this thesis focuses on domestic leisure destinations and does not include those zones in the model estimation. Figure 3 illustrates the NPVM traffic zones considered in the thesis. Each zone is associated with a unique identifier, the commune and canton to which it belongs, and its surface area in km^2.

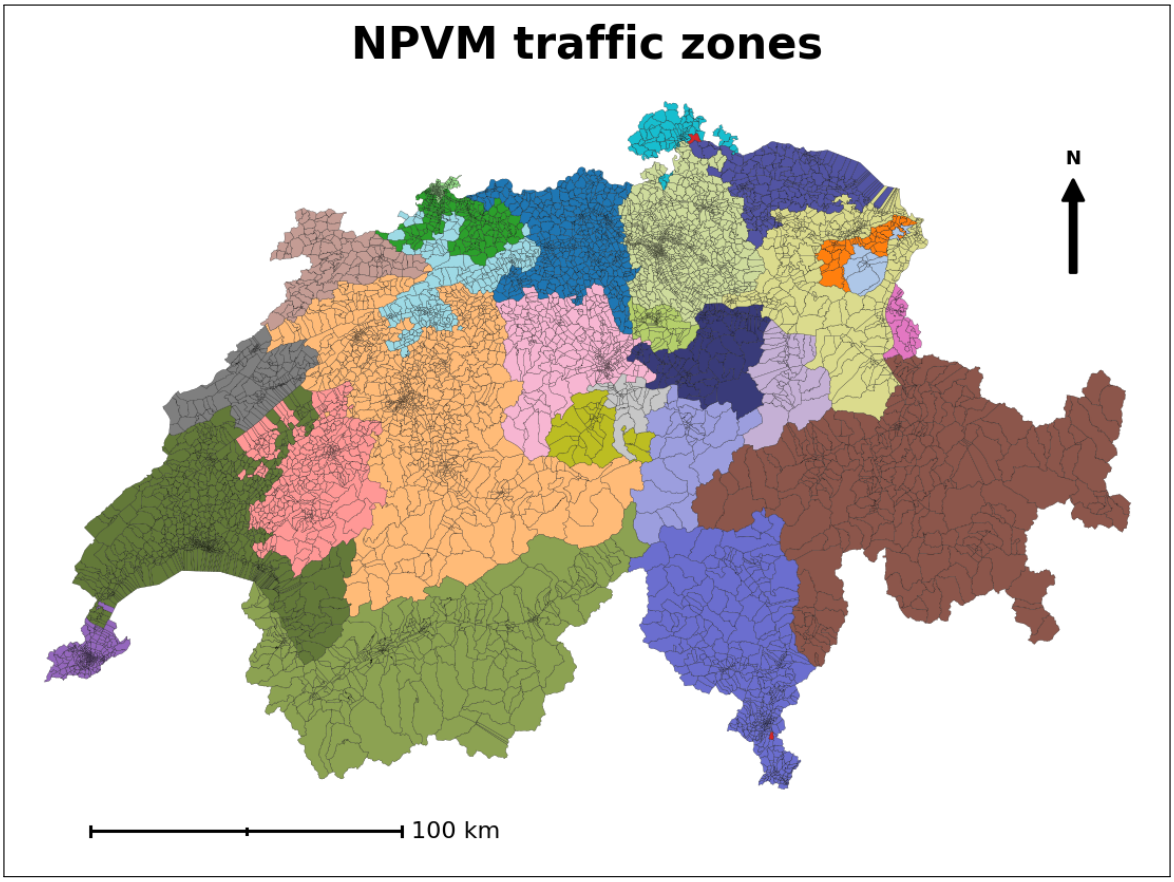


Figure 3: NPVM traffic zones 2017 in the study area, coloured by canton (VM-UVEK, 2017).

4.2 Observed leisure trips: Swiss Mobility and Transport Microcensus

The empirical basis for the demand-side analysis is the Swiss *Mobility and Transport Microcensus* (*Mikrozensus Mobilität und Verkehr*, MZMV/MTMC). The MTMC is Switzerland's national travel survey and is conducted approximately every five years by the

Federal Statistical Office (FSO/BFS) in collaboration with the Federal Office for Spatial Development (ARE) and a specialised survey institute (ARE & BFS, 2021, 2023). It combines socio-demographic and mobility-resource information with a one-day travel diary and additional modules covering less frequent and longer-distance travel.

In this thesis, I use respondents and trip records from both the 2015 and 2021 MTMC surveys. Pooling the two survey years increases the number of available observations and improves statistical power, which is particularly important in a destination-choice setting with a large alternative space and comparatively small subsamples. At the same time, pooling introduces a source of temporal heterogeneity: mobility behaviour and the territorial supply of opportunities may have changed between 2015 and 2021, and the 2021 survey was conducted in a context potentially affected by COVID-19. This issue is explored in the robustness analysis by retaining the survey year and testing sensitivity across the two survey waves (Section 5.5).

Survey design and data structure

The MTMC database is organised hierarchically, linking households, persons, and mobility records through unique identifiers. At the household level, the survey records household characteristics and mobility resources, while at the person-level it contains socio-demographic information. For each sampled household, one *target person* is selected as the reference person for the mobility survey.

For the target person, the core travel-diary component records all trips undertaken on the assigned reference day, including trip purpose, main mode, timing, distance, and geocoded origin and destination locations. For the purposes of this thesis, a particularly useful component is the domestic trip layer (*WegeInland*), which contains trips within the thesis study area and therefore allows a direct and consistent assignment to NPVM traffic zones. In addition, the MTMC includes supplementary modules covering less frequent travel. Of central importance here is the day-trip module (*Tagesreisen*), which is administered to a subset of respondents (approximately one third) and captures trips without overnight stay and with a minimum duration of 3 hours over a 14-day reference period. Only day trips whose geocoded origin and destination can be matched to the NPVM zone system retained in this thesis are retained. A separate module covers trips with overnight stays, but it is not used in this thesis.

The MTMC also provides a survey weight for each target person to account for the sampling design and for systematic non-response, so that weighted analyses better reflect the resident population. In this thesis, the person-level weight is attached to each trip record of the target person and used in model estimation. In practical terms, this means that each observed trip contributes in proportion to how strongly the corresponding person represents the underlying population, rather than counting all observations equally.

Operationalisation for this thesis

The 2015 MTMC includes 57,090 target persons, while the 2021 MTMC includes 55,018. Short leisure mobility is extracted from the domestic reference-day trip records (*WegeInland*) by retaining trips whose main purpose at destination is classified as leisure. Long leisure mobility is operationalised using the day-trip module (*Tagesreisen*) by retaining day trips whose reported purpose falls into leisure-related categories (see Appendix A.1). The short/long distinction is therefore defined by the survey component from which the observation is drawn. The age of the target person is then used to assign each observation to one of two age groups (6–64 and 65+), thereby completing the segmentation into four behavioural segments. For each retained leisure trip, the extracted attributes include travel distance (km), travel duration (minutes), person weight, and survey year. Using the origin and destination coordinates, each trip is assigned to NPVM traffic zones at origin and destination via spatial overlay. For exploratory and diagnostic purposes only, trips are additionally assigned to broad leisure activity types based on the reported detailed leisure-purpose categories. The preprocessing yields four segment-specific trip tables with origin and destination zone identifiers:

- **Young Short:** $N = 49{,}829$
- **Old Short:** $N = 13{,}648$
- **Young Long:** $N = 8{,}882$
- **Old Long:** $N = 1{,}988$

Each table contains one row per observed trip and includes the attributes listed in Table 4.

Table 4: Schema of the extracted leisure trip tables.

Attribute name	Description	Type
`HHNR`	Household identifier of the reporting target person	INT
`orig_zone`	NPVM traffic zone identifier of the trip origin	INT
`dest_zone`	NPVM traffic zone identifier of the trip destination	INT
`survey`	Survey year (2015 or 2021)	INT
`leisure_cat`	Broad leisure activity type (exploratory)	STRING
`dist_km`	Reported travel distance in kilometres	FLOAT
`dur_min`	Reported travel duration in minutes	INT
`WP`	Person weight of the reporting target person	FLOAT

Data exploration of leisure trips

To provide a compact descriptive view of the extracted leisure trips, observations are grouped into six broad leisure activity categories aligned with the domain-specific features discussed in Section 3.2: *Gastronomy*, *Visits (friends and relatives)*, *Outdoor*, *Culture*, *Sport*, and *Others*. Table 5 reports the weighted composition of observed leisure trips by segment and broad activity type. Shares are computed using the person weight `WP`. For short leisure trips from *WegeInland*, observations classified as *Not known* are excluded from the denominator when computing shares, although they remain valid observations for destination-choice estimation (76 for Young Short and 24 for Old Short).

Table 5: Weighted share (%) of observed leisure trips by broad activity type and segment.

Segment	Gastronomy	Visits	Outdoor	Culture	Sport	Others
Young Short	22.7	25.8	15.0	6.0	15.8	14.6
Old Short	25.1	24.1	17.9	5.6	7.6	19.7
Young Long	3.2	35.6	16.4	9.8	11.8	23.2
Old Long	4.4	32.8	24.2	8.0	3.7	27.0

The descriptive picture suggests some differences across both age groups and trip types. Older adults show a higher share of outdoor activities, gastronomy, and "Others", and a lower share of sport than younger individuals. For long leisure trips, visits become the dominant category in both age groups, while the relative importance of gastronomy decreases sharply. Overall, the contrast between short and long trips appears stronger than the contrast between age groups. Appendix A provides a more detailed exploratory analysis, extending this descriptive view to the spatial distribution of observed destinations. In particular, it documents how the geography of chosen leisure destinations varies across segments, across broad leisure categories, and in terms of local clustering patterns. These additional diagnostics show that heterogeneity also concerns where leisure activities tend to take place. Taken together, the exploratory evidence suggests that leisure behaviour may differ across the four segments. This, in turn, motivates the subsequent analysis of whether part of these differences can be explained by the territorial distribution of leisure opportunities, environmental qualities, and accessibility.

4.3 Destination attractivity: spatial indicators

The construction of the destination-side features $F_h(j)$ relies on a combination of official Swiss statistical datasets, authoritative topographic and environmental information, and globally available volunteered geographic information. In particular, OpenStreetMap (OSM) is used as the primary source for facility- and POI-based indicators.

A central principle in the construction of the supply-side indicators is to rely on OSM wherever possible in order to maximise the transferability of the proposed methodology across contexts. Comparable OSM data exist for most countries and regions, whereas official registers and administrative datasets often differ in definitions, spatial resolution, update frequency, and public accessibility. This consideration is particularly relevant for Liechtenstein and the two foreign enclaves included in the study area, for which Swiss governmental datasets do not always provide complete and harmonised coverage.

At the same time, relying on volunteered geographic information entails an important trade-off: no central authority guarantees the completeness, quality, or positional and thematic accuracy of the data. Coverage and attribute completeness may vary across space and feature types, so OSM-based indicators may be less reliable than those derived from official inventories. For this reason, whenever OSM-based indicators are used, their internal consistency is assessed, where possible, by comparison with official Swiss datasets. When OSM coverage is clearly insufficient, or when key attributes required for a given feature are not reliably available, official datasets are preferred. For Liechtenstein and the two enclaves, where Swiss federal statistics remain incomplete, missing values are complemented with external official sources or, where necessary, imputed. Overall, this combined use of volunteered and official data seeks to balance transferability, robustness, and data quality.

Data sources

The datasets used to construct the supply-side indicators are summarised below. Besides OSM, the official datasets used in this thesis are primarily based on administrative registers and official topographic or environmental models.

OpenStreetMap (OSM) is a global, openly licensed geospatial database maintained by an international community of contributors (OpenStreetMap contributors, 2026). Its fine-grained and continuously updated content makes it particularly suitable for representing heterogeneous leisure amenities and points of interest. OSM extraction is performed via Overpass Turbo (Overpass API) using reproducible, tag-based queries (Overpass Turbo contributors, 2026). The complete set of queries is reported in Appendix B.

Population and households (STATPOP 2024) is the official annual population and household statistics of Switzerland, based on administrative registers (BFS, 2024). It provides resident population and household information on a $100 \times 100\,\mathrm{m}$ grid. For data-protection reasons, small counts are partly aggregated. The dataset also includes breakdowns by age group, gender, and nationality.

Employment and workplaces (STATENT 2023) describes employment and workplaces at the level of individual business locations (*Arbeitsstätten*) (BFS, 2023). It covers all economic sectors and reports establishments, employees, and full-time equivalents, with breakdowns by sector. The data are geocoded on a $100 \times 100\,\mathrm{m}$ grid.

Land use and land cover (Arealstatistik 2025) provides nationwide information on land use and land cover at a resolution of $100 \times 100\,\mathrm{m}$ (BFS, 2025b). Based on manual interpretation of aerial imagery combined with topographic reference data, it distinguishes detailed classes of land use and land cover.

Topographic landscape model (swissTLM3D 2025) is the official high-resolution topographic vector model of Switzerland, maintained by the Federal Office of Topography

(Swisstopo, 2025). It provides a detailed representation of the natural and built environment, including transport infrastructure, hydrography, land cover, and leisure facilities. Additional official route data for snowshoe trails (Swisstopo, 2024a) and ski routes (Swisstopo, 2024b) are also used where relevant.

Glacier inventory (GLAMOS 2016) provides outlines and attributes of glaciers in Switzerland (GLAMOS, 2016).

Protected areas (2023/2025) are represented by combining official inventories and NGO reserve data. National parks and recognised regional or nature parks are taken from the Federal Office for the Environment inventory (BAFU, 2025). Nature reserves secured by Pro Natura are taken from the Pro Natura dataset (Pro Natura, 2023).

Operationalisation of spatial features

The feature blocks $F_h(j)$, corresponding to F_1–F_{10}, are constructed using a consistent workflow that is applied, with minor adaptations, across all leisure dimensions. For each block, candidate data sources are first selected based on conceptual relevance and data availability. For OSM, relevant features are identified through a targeted exploration of the tag space (OpenStreetMap Wiki contributors, 2025).

For each source layer, zone-level indicators are then obtained by aggregating the selected spatial features to NPVM traffic zones. These primary indicators include: (i) *counts* of discrete amenities, facilities, and other points of interest; (ii) *total lengths* of linear infrastructure; (iii) *total areas* of areal features; and (iv) additional *environmental descriptors* derived from land-cover and topographic data. Based on these primary aggregates, the workflow also generates derived indicators. In particular, selected variables are transformed into *density measures* by normalising them by zone area. Moreover, *diversity measures* are constructed by counting the number of distinct feature types present within a given feature block in each traffic zone.

Particular attention is paid to avoiding double counting. This issue arises within OSM when the same physical object may be selected through different keys or tagging combinations within the same feature block. The first control therefore consists of identifying and removing such duplicates whenever multiple tags point to the same entity. As an additional safeguard, a spatial deduplication rule is applied. If two OSM features have the same tag, the same name, and are located within 100 m of each other, only one is retained. Final counts of the selected OSM amenities are reported in Appendix C.

Where possible, the analysis also explored whether OSM features could be complemented by approximate *area-based* indicators, with the aim of capturing a rough proxy for their physical size or total capacity. In practice, this proved difficult because, for many OSM extractions, only a minority of the selected features are mapped as polygons. To address this, I experimented with the following procedure. Using the official swissTLM3D topographic dataset, OSM point features were first assigned to buildings through a spatial join whenever they fell directly within a building polygon. For points that did not intersect any building, the nearest building within at most 50 metres was assigned instead. The footprint area of the matched building was then used as the approximate area of the corresponding OSM entity. For the remaining point features still lacking an area, values were imputed by drawing from a distribution fitted on the observed areas available for the same amenity type. Since this procedure introduces substantial uncertainty, area-based OSM indicators are not retained in the final model specification, even though their feasibility was systematically explored.

Whenever possible, I also compared OSM-derived indicators with corresponding counts or aggregates derived from authoritative datasets. Depending on the case, this meant comparing OSM counts to official workplace counts from STATENT, OSM-based area proxies to official full-time equivalents (FTE), or OSM feature counts to objects extracted from swissTLM3D. In each case, I examined the correlation between the OSM-derived and official indicators across all traffic zones in order to obtain an empirical plausibility check on the internal consistency and quality of the OSM-based representation. A summary of these plausibility checks is reported in Appendix D.

Importantly, the purpose of this stage of the analysis is to construct a rich set of candidate indicators for each feature block, including alternative representations and, where available, official counterparts, from which a compact but sufficiently complete set can later be selected for model estimation. The final feature selection is documented in Section 5.1. To facilitate transparency and reproducibility, all traffic-zone variables produced during the construction workflow are stored directly as attributes in the traffic-zone dataset. A complete inventory of the resulting attributes is reported in Appendix E. The construction of the variables is described below. Note that for *F11: Spillovers (spatial autocorrelation)*, no direct extraction of spatial features is involved. This component is instead considered at the modelling stage (see Section 5.5).

F1: Gastronomy. The gastronomy block is constructed from OpenStreetMap (OSM) amenities related to food and drink consumption. The selected OSM features comprise restaurants and food courts (grouped into one category), bars, pubs and beer gardens (grouped into one category), cafés, fast-food outlets, and ice-cream shops. This yields a total of five distinct gastronomy amenity categories. From these features, three zone-level indicators are derived: `F1_gastr_count`, measuring the total number of gastronomy POIs in zone j; `F1_gastr_div`, measuring gastronomy diversity as the number of distinct categories present in the zone (range 0–5); and `F1_gastr_area_total`, measuring the total approximated footprint area of gastronomy-related amenities within the zone.

F2: Visits and social activities. In each NPVM traffic zone j, the *resident base* is operationalised as a proxy for the likelihood that relatives or friends live in a given zone. Two zonal indicators are constructed: total resident population and total private households. For Switzerland, both indicators are derived from STATPOP 2024. STATPOP grid-cell values are spatially assigned to the traffic zones and aggregated to obtain `F2_pop_total` (total resident population in zone j) and `F2_hh_total` (total private households in zone j). For Liechtenstein and the two foreign enclaves, resident population totals are compiled from external official demographic sources (Amt für Statistik Liechtenstein, 2025; Gwind S.r.l., 2023; UrbiStat S.r.l., 2026). Household totals are imputed by applying the Swiss average persons-per-household ratio implied by STATPOP to the externally sourced population totals and rounding the resulting values to the nearest integer.

F3: Outdoor recreation. This block captures the potential of a traffic zone to support leisure activities in open and natural environments. Since outdoor recreation depends on the joint presence of multiple landscape elements rather than on a single facility type, $F_3(j)$ is operationalised through a set of complementary indicators.

Everyday walking opportunities are captured by `F3_outdoor_walk`, defined as the total length (km) of officially designated hiking trails (*Wanderwege*) from swissTLM3D within each zone, excluding mountain and alpine trail classes (*Bergwanderweg*, *Alpinwanderweg*). Protected environments are represented by `F3_outdoor_protected` (binary presence of any protected area) and `F3_outdoor_protected_area` (protected area in km^2), based on the union of BAFU parks and Pro Natura reserves (N = 819).

Water-related features are captured separately for lakes and rivers. Lakes are derived from swissTLM3D hydrography (N = 1,272), since OSM water polygons include many very small or artificial water bodies that are not necessarily recreationally relevant. Three lake indicators are derived: `F3_outdoor_lake_shore` (total shoreline length per zone), `F3_outdoor_lake_bin` (binary presence of shoreline), and `F3_lake_area` (intersected lake surface). Rivers, streams and canals are derived from OSM and aggregated as `F3_outdoor_river` (total length in km). To avoid artefacts and double counting, river segments falling inside reconstructed lake polygons are removed.

Hard outdoor POIs capture a compact set of discrete features that signal more "demanding" recreation opportunities. In this study, these features are represented through six categories derived from OSM: aerialways, huts, glaciers, caves, waterfalls, and viewpoints. Two zone-level indicators are constructed: `F3_outdoor_hard_count`, measuring the total number of such POIs in zone j, and `F3_outdoor_hard_diversity`, measuring the number of distinct hard-outdoor categories present in the zone (range 0–6).

Soft outdoor opportunities capture low-intensity, broadly accessible facilities that support everyday recreation. A total of eight categories are extracted from OSM: playgrounds, picnic sites and picnic tables (treated as one category), parks, dog parks, firepits, bird hides, beach resorts, and marinas. From these features, two zone-level indicators are constructed: `F3_outdoor_soft_count`, measuring the total number of soft outdoor POIs in zone j, and `F3_outdoor_soft_diversity` (range 0–8). The analysis also explored whether their spatial extent could be approximated through area-based indicators, using both OSM polygons only and, where possible, area associations with corresponding areal objects from swissTLM3D. In practice, this proved difficult, and the count- and diversity-based indicators are therefore retained as the main operationalisation.

A land-use mix indicator is used as a compact descriptor of environmental heterogeneity. It is derived from *Arealstatistik* 2025 using the 10-class land-use aggregation `NOLU04_10`, as reported in Table 6 (BFS, 2019). Grid cells are assigned to NPVM traffic zones by spatial intersection, and the indicator is then computed for each zone as the normalised Shannon entropy of the land-use composition:

$$F3\_outdoor\_LUmix(j) = \frac{-\sum_{k=1}^{K} p_{jk} \ln(p_{jk})}{\ln(K)}, \qquad K = 10, \tag{9}$$

where p_{jk} denotes the share of land-use class k in zone j. The resulting values lie in the interval $[0, 1]$, with higher values indicating a more heterogeneous land-use composition. To reduce artefacts in traffic zones bordering large lakes, the share of the water class is capped at 40% prior to the entropy calculation. For the two enclaves (`Enk`), values are imputed as the mean land-use mix of Swiss zones within a 20 km neighbourhood (nearest-zone fallback). For Liechtenstein, values are imputed using donor zones from neighbouring cantons (SG, AI, AR) in order to obtain plausible zone-level values.

Table 6: Arealstatistik land-use classes used for the land-use mix indicator.

Code	Class name	Code	Class name
100	Building areas	220	Arable and grassland
120	Transport surfaces	240	Alpine grazing areas
140	Special urban areas	300	Forest
160	Recreational areas and cemeteries	400	Lakes and rivers
200	Orchards/vineyards/horticulture	420	Unproductive land

F4: Cultural opportunities. The cultural block is constructed from OpenStreetMap (OSM) amenities and tourism-related features representing cultural and entertainment opportunities. The selected OSM features comprise museums, libraries, theatres, galleries, cinemas, event venues, arts centres, music venues, zoos, conference centres, exhibition centres, planetariums, and aquariums, yielding a total of 13 distinct cultural categories. From these features, three zone-level indicators are derived: `F4_cult_count` (total number of cultural POIs in zone j); `F4_cult_diversity` (range 0–13); and `F4_cult_area_total` (approximated footprint area).

F5: Sport opportunities. This block captures the availability of amenities and opportunities for *active sport*. The selected OSM sport-related features comprise pitches, fitness centres, sports centres, running and sports tracks, slipways, horse-riding facilities, miniature golf, swimming areas, sports halls, pavilions, dance facilities, ice rinks, stadiums, riding halls, and disc golf courses, yielding a total of 15 distinct sport categories. From these features, three zone-level indicators are derived: `F5_sport_count`, `F5_sport_div` (range 0–15), and, where feasible, `F5_sport_area_total`. The latter was approximated by transferring areas from official swissTLM3D sport polygons to nearby OSM POIs.

Outdoor sport routes are represented separately as linear infrastructure. Route information is compiled from official datasets for hiking, ski, and snowshoe infrastructure, while via ferratas are extracted from OSM. In the case of hiking, the official swissTLM3D data distinguish three route categories, namely *Wanderweg* (walking trail), *Bergwanderweg* (mountain hiking trail), and *Alpinwanderweg* (alpine hiking trail). These linear layers are intersected with traffic-zone polygons and aggregated in kilometres as `F5_sport_out_length`. In addition, the route-type diversity measure `F5_sport_out_div` counts the number of distinct route types present in the zone (minimum length: 100 m), yielding values between 0 and 6.

F6: Other leisure domains. $F_6(j)$ captures a residual set of leisure-related facilities not covered by the previous features. The selected OSM features are grouped into five macro-categories: *nightlife*, including casinos and nightlife venues; *social*, including community and social centres; *spiritual*, including places of worship, cemeteries, and monasteries; *fun*, including gaming and amusement facilities such as arcades, escape rooms, water parks, and theme parks; and *wellness*, including massage facilities. The resulting indicators are the category-specific counts `F6_others_nightlife`, `F6_others_social`, `F6_others_spiritual`, `F6_others_fun`, and `F6_others_wellness`, together with a diversity indicator `F6_others_diversity`, defined as the number of macro-categories present in zone j (range 0–5). In addition, `F6_others_count` aggregates all categories except the spiritual category.

F7: Diversity indices. In addition to the domain-specific diversity measures introduced above, composite diversity indices are constructed to capture the overall breadth of leisure opportunities within each traffic zone. The inputs are the feature-level diversity indicators derived for the main dimensions. Each diversity input is also rescaled to the interval $[0, 1]$ using min–max normalisation. Several composite summaries are then computed in raw and normalised form. First, a generic diversity index combines all available diversity components into an overall measure of leisure variety within the zone. For this purpose, a raw sum (`F7_div_sum`) is retained together with its normalised counterpart, while the mean is stored only in normalised form (`F7_norm_mean`). Second, separate composite indices are constructed for an *urban* bundle (gastronomy, sport facilities, culture, and other facilities) and for an *outdoor/nature* bundle (outdoor hard features, outdoor soft features, and sport-route diversity). For both bundles, sum and mean versions are stored, again in raw and normalised form (e.g. `F7_urban_sum`, `F7_outdoor_mean_norm`).

F8: Density indicators. To ensure comparability across zones of very different size, density variants are created by normalising selected supply measures by zone area. For any zone-level variable $X(j)$, the raw density is defined as

$$F8_{X,\mathrm{dens}}(j) = \frac{X(j)}{\mathrm{Area}_{\mathrm{km}^2}(j)}. \tag{10}$$

Density variables are stored with the prefix `F8_` and the suffix `_dens`. In addition, four composite densities are computed by summing conceptually related quantities and dividing by area: `F8_POI_urban_dens` (gastronomy, sport facilities, culture, and others), `F8_POI_natural_dens` (outdoor hard and soft features), `F8_lengths_dens` (linear recreation infrastructure), and `F8_POI_total_dens` (urban and natural POIs). Densities are stored in units per km^2.

F9: Monetary cost. In the present study, a dedicated monetary-cost variable is not operationalised. Price data for leisure-related facilities are difficult to obtain in a reproducible way, as they would often require scraping platform-based information or using proprietary APIs. More general economic proxies are also problematic, since they are rarely available at the spatial granularity of NPVM traffic zones and are often difficult to access. For these reasons, no direct monetary-cost indicator is retained in the final specification.

F10: Supporting infrastructure and services. $F_{10}(j)$ captures a compact set of basic support amenities that can improve comfort and reduce friction during leisure trips, independently of the main activity type. The selected OSM features include drinking-water points, fuel stations, electric charging stations, public toilets, and ATMs. These amenities are represented as points and aggregated to traffic zones through a single indicator, `F10_superinfra`, defined as the total count of such support amenities in zone j. In contrast to other feature blocks, no diversity measure is constructed.

4.4 Origin–destination accessibility

This subsection defines the accessibility control used in the empirical specification. In travel modelling, impedance can be operationalised in multiple ways depending on data availability and modelling goals. A common classification distinguishes threshold-based (isochrone) measures, gravity-based (distance-decay) measures, and utility-based measures grounded in discrete choice theory (Dong et al., 2006). Utility-based accessibility interprets impedance through the expected maximum utility of a set of travel options (logsum) and has been used in empirical applications (Guzmán et al., 2023). Impedance can also be represented by simpler metrics such as beeline distance, network distance, car travel time, public-transport travel time, or other generalised costs. Since the main contribution of this thesis lies in the construction of destination-side attractivity, the proposed framework remains, in principle, compatible with different impedance specifications. In the empirical setup, this thesis adopts a utility-based measure consistent with the SIMBA MOBi framework, which also facilitates direct integration with the existing data pipeline. The sensitivity of the results to alternative impedance representations is assessed separately in Section 5.5.

The adopted measure is based on origin–destination (OD) skim matrices and zonal accessibility attributes available within the SIMBA data pipeline. A skim matrix is a precomputed OD table that summarises level-of-service (LOS) conditions between all zones (e.g. travel time, distance) and is used as a direct input to utility calculations. These inputs are already prepared for operational modelling, but their construction is transparent and can in principle be reconstructed from the underlying transport skims and zonal attributes. For each OD pair (i, j), systematic travel utilities are computed for the four available modes: walk, bike, car, and public transport (PT). These utilities depend on the conditions between origin and destination, including distance, travel time, access and egress times, service frequency, transfers, and selected zonal attributes such as car access times and parking costs. The full mode-specific utility equations and parameter values, adopted here as representative coefficients inherited directly from the SIMBA MOBi framework without recalibration, are reported in Appendix F.

The multimodal accessibility indicator is then defined as the expected multimodal utility (EMU) across the modes considered, i.e. the logsum:

$$\mathrm{EMU}_{ij} = \ln \sum_{m \in \mathcal{M}} \exp\left(\frac{U_{ijm}}{\theta_{\mathrm{mode}}}\right), \qquad \theta_{\mathrm{mode}} = 1. \tag{11}$$

where $\mathcal{M} = \{\mathrm{walk}, \mathrm{bike}, \mathrm{car}, \mathrm{PT}\}$ and θ_{mode} is the scale parameter of the logsum. In the present application, it is fixed to 1. Higher values indicate easier access, or equivalently lower impedance, from origin i to destination j. Missing or unreachable public-transport relations are treated as unavailable when forming the logsum, i.e. they contribute zero to the sum inside the logarithm.

The OD skim matrices required for this construction must cover the full thesis zone system, including Liechtenstein and the two foreign enclaves. The skim exports available for MOBi 4.0 (2017 delivery) do not include these areas consistently, which would create structurally missing OD relations. MOBi 5.0 skim matrices (2023 delivery) are therefore used, restricting them to the thesis zone set. Because these skims correspond to the 2023 data delivery, EMU variants are computed using both 2017 and 2023 zonal accessibility/cost attributes (e.g. car access times and parking costs). Figure 4 shows a destination-level aggregation of the utility-based accessibility measure across the thesis zone system. For each destination zone j, the mapped value corresponds to the sum of EMU_{ij} over all origin zones i, used for descriptive analysis only.

A potential concern is that the MTMC survey waves (2015/2021) reflect transport conditions that may differ slightly from the 2023 skim snapshot. To assess the sensitivity of this choice, I compute EMU variants using both 2017 and 2023 zonal accessibility/cost attributes while keeping the MOBi 5.0 skims fixed, and I compare the resulting measure with simpler alternative impedance proxies. EMU variants based on 2017 versus 2023 zonal attributes are almost perfectly correlated (Pearson $\approx$ 0.99 on N = 30,000 OD samples). At the trip level (all leisure trips; n = 74,347 with valid distances), EMU correlates strongly and negatively with reported trip distance (Pearson ≈ -0.91). EMU also correlates negatively with the beeline distance between traffic-zone centroids (Pearson ≈ -0.61 for car and Pearson ≈ -0.66 for PT). These two results indicate that the accessibility measure used here is closely aligned with simpler notions of impedance and distance, and therefore behaves as a sensible accessibility proxy.

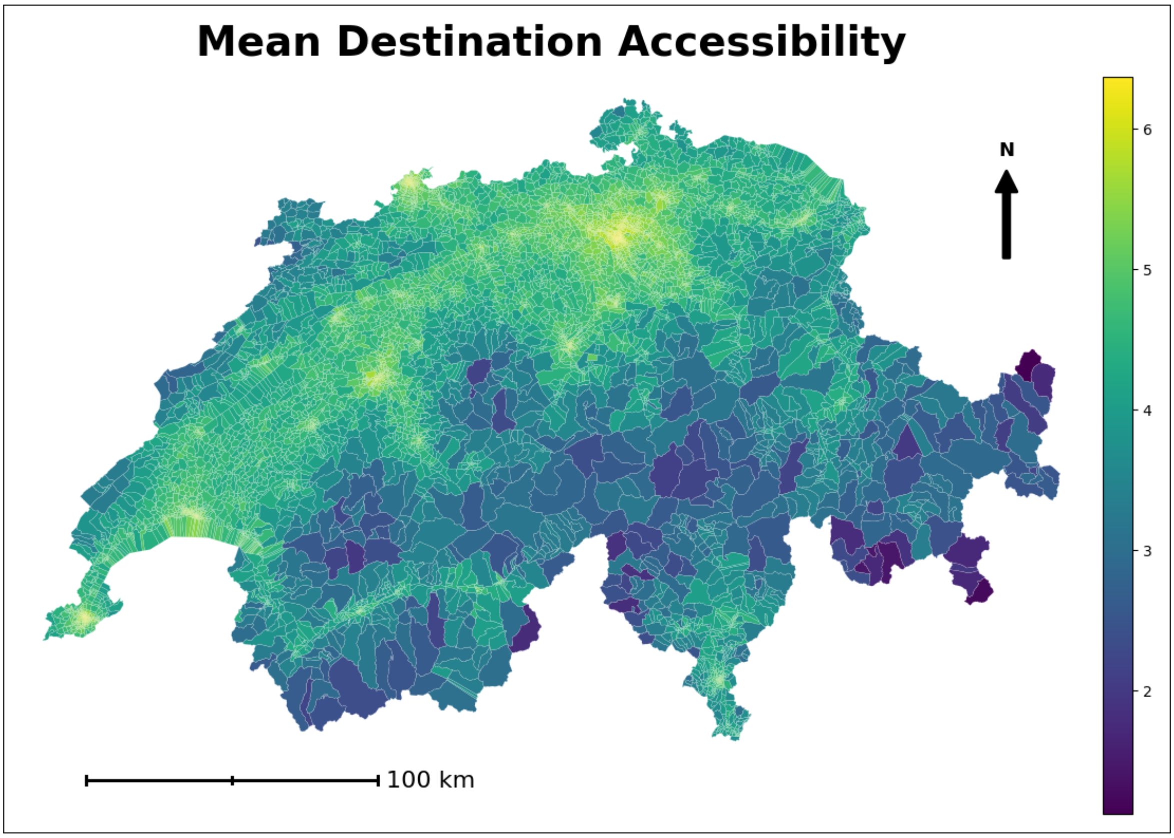


Figure 4: Destination-aggregated spatial pattern of the utility-based accessibility measure across NPVM traffic zones. Higher values indicate destinations that are, on average, more easily accessible within the multimodal transport system.

5 Methodology: specification, estimation and evaluation

This chapter presents the methodology used to estimate four segment-specific destination choice models for leisure trips at the NPVM traffic-zone level. A central element of the specification is a new destination-side attractivity term, $A_j(s)$, which is estimated separately for each behavioural segment. Section 5.1 describes how the broad pool of zonal indicators is reduced to a compact set of 13 destination features $F_h(j)$ entering the attractivity specification. Section 5.2 then presents the two main experimental design choices required for estimation at the national scale: the train–test split of observations and the sampling of non-chosen destination alternatives. Section 5.3 introduces the estimation and evaluation strategy, together with the benchmarking framework used to compare the proposed model against alternative specifications. Section 5.4 explains how the estimated coefficients are interpreted and compared across segments. Section 5.5 outlines the robustness analyses used to assess the sensitivity of the results to key modelling choices. Finally, Section 5.6 describes how the estimated attractivity term is operationalised, integrated into SIMBA MOBi, and evaluated within the simulation framework.

5.1 Attractivity index construction

This section describes how the broad pool of destination-side indicators introduced in Chapter 4 and listed in Appendix E is reduced to a compact set of variables entering the attractivity specification. The starting database contains multiple alternative operationalisations of similar underlying constructs (e.g. counts vs. densities, composite vs. specific measures, or OSM-based proxies vs. official counterparts). Including all candidates jointly would inflate dimensionality, reduce interpretability, and increase the risk of unstable estimates. I therefore adopt a staged selection strategy, with the goal of converging to a compact and interpretable specification of destination-side indicators.

The selection is guided by four main rationales. First, *data quality*: wherever possible, indicators should rely on sources that are sufficiently complete and internally consistent, with particular attention to OpenStreetMap-based measures (Appendix D). Second, *interpretability*: each retained indicator should have a clear semantic meaning and correspond to a recognisable dimension of leisure opportunities. Third, *limited semantic redundancy*: the retained indicators should capture distinct aspects of leisure attractivity. For this reason, the selection follows the conceptual feature structure introduced in Section 3.2 and aims at a reasonably balanced coverage of the main leisure dimensions. Fourth, *collinearity screening*: because many amenities and opportunities co-locate in the same zones, semantic overlap can also manifest itself numerically through strong pairwise correlations and multicollinearity. I therefore use pairwise correlations and variance inflation factors (VIF) as diagnostic tools to identify potential redundancy. VIF measures how strongly the variance of a coefficient is inflated by linear dependence with the other regressors

(Wooldridge, 2016). In this thesis, it is used only as a diagnostic aid, since commonly used thresholds such as 5 or 10 remain rules of thumb and depend on the context (O'Brien, 2007).

Starting from the full set of 133 destination-side indicators, I apply a stepwise selection. The decisions are summarised below (N is the number of variables removed at each step).

- **Official counterparts replaced by OSM-based proxies** ($N = 15$). Whenever an OpenStreetMap-based indicator was judged sufficiently plausible and internally consistent, I excluded the corresponding official count or employment-based proxy.
- **Area-based POI proxies removed** ($N = 11$). I excluded the area- and footprint-based approximations of POI size or capacity because their quality is substantially weaker than that of count-based indicators.
- **Subtype counts removed** ($N = 41$). I excluded counts for specific subcategories within broader leisure domains (e.g. museums, theatres, individual "others" categories such as *social*, *fun*, *wellness*, and *nightlife*). These detailed indicators create unnecessary redundancy.
- **Redundant features removed** ($N = 10$). I excluded five small groups of variables because their information content was judged to overlap with other retained indicators that are more directly interpretable. First, I removed binary presence indicators, as they provided limited additional information relative to their metric counterparts. Second, for linear infrastructure, I retained total length rather than density, since total supply was judged more directly interpretable than length per km^2. Third, I retained resident population and excluded households, as population was considered the more plausible proxy for social visits in this context. Fourth, I removed protected-area indicators because the outdoor dimension is already represented through a rich set of other variables. Fifth, I excluded `F3_outdoor_walk` as a separate indicator because its information is already captured by `F5_sport_out_length`.
- **Within-feature diversity measures removed** ($N = 14$). The first VIF analysis showed that the feature-specific diversity indicators are highly collinear with the aggregated diversity measures. I therefore excluded all diversity measures at the single-feature level, both normalised and non-normalised.
- **One proxy for overall diversity** ($N = 12$). I retain only `F7_urban_sum`, i.e. the sum of the raw diversity counts across urban leisure domains. This choice reflects the aim of capturing the overall breadth of urban leisure opportunities through a single composite indicator. I retain only the urban bundle because the environmental heterogeneity of destinations is already captured by the land-use mix variable. Among the aggregate diversity candidates, `F7_urban_sum` showed the lowest VIF.

- **Blue-space representation simplified** ($N = 5$). I simplified the representation of blue-space opportunities by removing river-based indicators. River indicators were excluded because the OSM extraction includes heterogeneous water features, including minor channels that are not necessarily meaningful as leisure opportunities. For the lake dimension, I retained only shoreline-length density (`F8_outdoor_lake_shore_dens`). This choice is motivated by two reasons. First, lake area can be artificially inflated for traffic zones extending into large transboundary lakes. Second, raw shoreline length tends to over-reward Alpine zones with many small lakes relative to major lakeside urban destinations.

- **Spiritual indicator excluded** ($N = 2$). The exclusion of the spiritual component is one of the more delicate decisions. Empirically, the corresponding trip shares are relatively small (around 4% for older adults and 1% for younger individuals). Conceptually, visits to places of worship are also likely to be driven by specific motives and local familiarity, often favouring the nearest location rather than a broad competitive set of destinations.

- **One generic urban-density control** ($N = 10$). The final major decision concerns whether POI-based features should enter the model as absolute counts or as area-normalised densities. I retain raw counts as the default specification. This choice is supported by the fact that NPVM traffic zones are already designed to be relatively comparable in structural terms. Moreover, many density measures are strongly correlated with each other. I retain a single generic urban-density control, `F8_POI_urban_dens`, constructed from the main urban leisure domains. I focus this density control on urban amenities because the outdoor and nature dimensions are already represented by several dedicated indicators. Density-based alternatives to raw counts are then examined explicitly in the robustness analysis (Section 5.5).

After this staged reduction, the remaining set contains 13 destination-side variables that are, as far as possible, semantically distinct, transparent in interpretation, and jointly representative of the main leisure domains. As a first preprocessing step, I apply a $\log(1 + x)$ transformation to all retained indicators except the land-use mix variable `F3_outdoor_LUmix`, which is already bounded in the interval $[0, 1]$. This is motivated by the fact that the zonal variables are non-negative and often strongly right-skewed. In particular, some structurally atypical zones in Liechtenstein, where the spatial units correspond to municipalities rather than subdivided traffic zones, may exhibit inflated values relative to the rest of the study area. The $\log(1 + x)$ transformation reduces the influence of such extreme values while preserving zero observations and the rank ordering of zones (West, 2022; Wooldridge, 2016). Maps of the transformed zonal indicators are reported in Appendix G. Table 7 summarises the retained destination-side indicators and their corresponding VIF values, whose levels are moderate overall. The second preprocessing step, namely z-standardisation, is applied immediately before model estimation and is described in Section 5.3.

Table 7: Final set of retained zonal indicators for the attractivity specification and associated VIF values. Numbering is kept consistent with the coefficient tables reported later in the thesis.

Block	Retained indicator	Name	VIF
F1	`F1_gastr_count`	1. Gastronomy POIs	3.21
F2	`F2_pop_total`	2. Resident population	1.26
F3	`F8_outdoor_lake_shore_dens`	3. Lake shore density	1.18
F3	`F3_outdoor_hard_count`	4. Hard outdoor POIs	2.05
F3	`F3_outdoor_soft_count`	5. Soft outdoor POIs	1.59
F3/F7	`F3_outdoor_LUmix`	6. Land-use mix	1.41
F4	`F4_cult_count`	7. Cultural POIs	1.55
F5	`F5_sport_count`	8. Sport POIs	2.31
F5/F3	`F5_sport_out_length`	9. Outdoor sport route length	4.00
F6	`F6_others_count`	10. Other leisure POIs	1.24
F7	`F7_urban_sum`	11. Urban diversity index	4.59
F8	`F8_POI_urban_dens`	12. Urban POIs density	4.22
F10	`F10_superinfra`	13. Support/service amenities	2.92

The destination-side attractivity term is modelled as a linear index of the selected zonal indicators. Let $\widetilde{F}_h(j)$ denote the final transformed version of indicator h for zone j, after applying $\log(1+x)$ and subsequent z-standardisation. The coefficients $\beta_h^{(s)}$ are estimated separately for each segment s.

$$A_j(s) \;=\; \sum_{h=1}^{H} \beta_h^{(s)}\, \widetilde{F}_h(j), \qquad H = 13, \tag{12}$$

5.2 Experimental design decisions

This subsection outlines the empirical design adopted in the thesis. It first presents the out-of-sample split used to assess predictive performance and transferability, and then describes the sampled choice-set construction required in the presence of a very large destination universe.

Train/out-of-sample split

The out-of-sample (OOS) split is used to assess predictive performance on observations not employed for model estimation, since in-sample fit is typically optimistic and may overstate generalisation ability. The validation literature therefore recommends evaluating discrete-choice models on held-out data rather than relying exclusively on in-sample goodness-of-fit measures (Parady et al., 2021; Parmar & Delle Site, 2025).

In this thesis, the default OOS design is a fixed 80/20 holdout defined at the person level within each segment s. All trips associated with the same respondent are assigned either to the training set or to the test set, so that repeated behaviour of the same individual cannot appear on both sides of the split. More generally, the splitting scheme should be aligned with the dependence structure of the data and with the intended prediction target, rather than treating all observations as fully independent (Yates et al., 2023). A possible alternative would be a more demanding split defined by leaving origins out, so that origins appearing in the test set are absent from training. By analogy with leave-one-group-out validation, this would likely provide a stricter assessment of performance when predicting to new spatial contexts.

Because estimation relies on survey expansion weights (WP), the 80/20 split is defined in weighted terms rather than by raw observation counts. Within each segment, respondents are first randomly ordered and then assigned to the test set until the cumulative sum of their survey weights reaches approximately 20% of the total segment weight; the remaining respondents are assigned to training. To avoid having the OOS sample dominated by a single MTMC wave, the holdout is also balanced across 2015 and 2021 in weighted terms, so that both waves contribute comparably to the evaluation set. This makes the reported OOS performance less dependent on the composition of one single survey wave. The split is generated once per segment using a fixed random seed and then kept unchanged throughout the analysis.

Sampled choice-set construction

The destination-choice problem in this thesis involves a very large universal choice set, since in principle each of the $J = 7{,}978$ NPVM traffic zones can serve as a leisure destination. Estimating a multinomial logit (MNL) model on the full destination universe for every observed trip would therefore be computationally burdensome. A standard solution in the discrete-choice literature is to estimate the model on *sampled choice sets* that include the chosen destination together with a subset of non-chosen alternatives (Dekker et al., 2025; Guevara & Ben-Akiva, 2013; Train, 2009). Within this framework, two design choices become central: how the non-chosen alternatives are sampled from the full destination universe, and how large the sampled choice set should be. These choices can affect both estimation accuracy and predictive performance (Becker et al., 2025; Dekker et al., 2025).

In this thesis, the main sampling design is *uniform random sampling of alternatives*. For each observed trip, the sampled choice set contains the chosen destination and K non-chosen destinations drawn uniformly at random from the full destination universe. This design is adopted because it is simple, transparent, and straightforward to implement and interpret. At the same time, uniform sampling has a practical limitation in spatial destination-choice settings. When the universal choice set is very large, many uniformly drawn non-chosen destinations may be formally feasible but behaviourally less relevant, for example because they imply very low accessibility from the given origin (Becker et al., 2025).

An alternative is to sample non-chosen destinations in a more targeted way, using information on their behavioural plausibility, so that more relevant competitors are more likely to enter the sampled set (Becker et al., 2025). In the present application, one possible implementation would be to inform the sampling by the accessibility values EMU_{ij} observed in the trip data. More specifically, for a given origin zone i, higher inclusion probabilities can be assigned to non-chosen destinations j whose EMU_{ij} values are more consistent with the range of accessibility observed for realised trips. In this sense, the sampling rule would favour destinations that are more plausible competitors conditional on the origin. Appendix H reports the empirical distribution of chosen-destination EMU_{ij} values by segment in order to illustrate the accessibility ranges underlying this idea.

However, this second strategy comes at a methodological cost. When alternatives are sampled with unequal probabilities, the sampling design must be accounted for directly in the systematic utility estimated by the model. In the logit case, this is done through the McFadden correction, that is, by adding a term based on the logarithm of the sampling probability to the utility specification alongside accessibility and attractivity (Dekker et al., 2025; McFadden, 1978). The purpose of this correction is to adjust for the fact that some alternatives are made more likely than others to enter the sampled choice set because of the sampling rule itself. This is not necessary under uniform random sampling because this design satisfies the *uniform conditioning* requirement: the probability of generating a sampled set does not depend on which alternative in that set is the realised choice (Train, 2009). For this reason, uniform sampling is used here as the main specification, while accessibility-informed sampling is treated only as a robustness analysis check.

Finally, the size of the sampled choice set must also be determined. Before proceeding to model estimation and evaluation, this is addressed in a separate model-selection exercise, the results of which are reported in Appendix H. Starting from the outer 80% training sample, an additional weighted 80/20 inner split is created within the training data for model-selection purposes. This inner split is defined at the observation level and is used only to compare alternative values of K. Different values of K are then compared on this inner validation set, using uniformly sampled non-chosen alternatives, and evaluated with the same performance metrics and full-choice-set diagnostics later employed for the main model assessment (see Section 5.3). Across segments, performance on the inner validation set generally improves as K increases but tends to stabilise around $K = 1000$, while larger sampled choice sets yield only limited additional gains relative to their computational cost. On this basis, the baseline specification retains $K = 1000$ non-chosen alternatives per observation.

5.3 Model estimation and evaluation metrics

This subsection presents the modelling framework used to estimate destination choice. It introduces the multinomial logit specification adopted in this thesis and explains how its main components, defined in Chapter 4, are preprocessed and incorporated into the estimation procedure. Next, it describes how predictive performance is assessed on out-of-sample data, with particular attention to the evaluation metrics used to assess model performance. Finally, it outlines the benchmarking strategy through which the proposed utility specification is compared against alternative formulations.

The empirical analysis is conducted separately for the four behavioural segments considered in this thesis, denoted by $s \in \{\text{Young–Short}, \text{Old–Short}, \text{Young–Long}, \text{Old–Long}\}$. Rather than estimating a single pooled model with interaction terms, I estimate one destination-choice model for each segment s. This choice is motivated by the aim of obtaining segment-specific behavioural predictions and coefficients that remain directly interpretable (Train, 2009). For each observed trip n belonging to segment s, destination choice is modelled within the random-utility framework. Let i denote the origin zone and j a potential destination. Utility is written as

$$U_{nij}^{(s)} = V_{nij}^{(s)} + \varepsilon_{nij}^{(s)}, \tag{13}$$

where $\varepsilon_{nij}^{(s)}$ is an i.i.d. extreme-value error term. Under this assumption, destination choice follows the standard multinomial logit (MNL) form (Train, 2009). In the present case, the systematic component of utility for observation n can be written as:

$$V_{nij}^{(s)} \;=\; \alpha^{(s)}\,\mathrm{EMU}_{ij} \;+\; A_j(s), \tag{14}$$

where EMU_{ij} is the accessibility-based impedance measure (see Section 4.4) and $A_j(s)$ is the segment-specific leisure attractivity index defined in Equation 12.

The destination-side indicators entering $A_j(s)$ (apart from *Land-use mix (6)*) are first transformed using $\log(1 + x)$. All of them (including *Land-use mix (6)*) are then z-standardised, where μ_h and σ_h denote the mean and standard deviation of the transformed variable h, computed over the full universe of traffic zones.

$$\widetilde{F}_h(j) \;=\; \frac{\log(1 + F_h(j)) - \mu_h}{\sigma_h}. \tag{15}$$

This standardisation places all destination attributes on a common scale, so that each transformed variable has mean 0 and standard deviation 1. In practical terms, it improves numerical stability during estimation and makes comparisons of coefficient magnitudes more meaningful (Wooldridge, 2016). This transformation can yield negative values of the estimated utility. This is not problematic, because when the utility enters Equation 16, it is mapped to a positive probability.

For estimation, each observed trip is associated with a set composed of the chosen destination and $K = 1000$ sampled non-chosen alternatives. Let $\mathcal{K}_n$ denote the sampled choice set for observation n, including the observed destination. The probability that observation n chooses destination j within this sampled set is then given by

$$P_{nij}^{(s)} = \frac{\exp\left(V_{nij}^{(s)}\right)}{\sum_{k\in\mathcal{K}_n} \exp\left(V_{nik}^{(s)}\right)}, \tag{16}$$

Model parameters are estimated by weighted maximum likelihood using the MTMC weights WP_n. Equivalently, estimation minimises the weighted negative log-likelihood over all training observations in segment s, using the Adam gradient-based optimiser (50 epochs, batch size 2048, adaptive learning rate initialised at 0.03):

$$\mathcal{L}^{(s)}(\alpha^{(s)}, \beta^{(s)}) = -\sum_{n\in\mathcal{T}_{\text{train}}^{(s)}} WP_n \log P_{niy_n}^{(s)}, \tag{17}$$

where $\mathcal{T}_{\text{train}}^{(s)}$ denotes the set of training observations in segment s, and y_n denotes the observed destination chosen in observation n. In practical terms, the estimation procedure selects the values of $\alpha^{(s)}$ and $\beta^{(s)}$ that assign the highest possible probability to the destinations actually observed in the training data, while accounting for the survey weights. In summary, the baseline estimation setup consists of four segment-specific models defined over a destination universe of 7,978 traffic zones. Utility combines the EMU-based accessibility term with 13 destination-side variables. The retained destination variables are preprocessed through $\log(1+x)$ and then globally z-standardised. Estimation is based on the pooled 2015 and 2021 MTMC data, using a person-based weighted 80/20 train–test split and sampled choice sets with $K = 1000$ uniformly drawn non-chosen alternatives per observation. The baseline specification includes neither regularisation nor an explicit spatial-autocorrelation component.

For each segment s, the model is estimated on the corresponding training sample only. The estimated coefficients are then kept fixed and applied, without further re-estimation or recalibration, to the corresponding out-of-sample (OOS) observations of the same segment. For each OOS observation, predicted choice probabilities are computed over the full destination universe $j = 1, \ldots, J = 7{,}978$. This ensures that model performance is assessed under the full choice environment faced by individuals. It also allows aggregate predicted destination shares to be derived consistently over the whole study area, which is useful for evaluating whether the model reproduces plausible spatial attraction patterns.

Model performance on the OOS sample is then assessed using the following evaluation metrics.

- **Weighted average negative log-likelihood (lower is better).** On the OOS sample, this is defined as the weighted mean negative log-probability assigned to the observed destination:

$$\overline{\mathrm{NLL}}_{\mathrm{OOS}} = -\frac{\sum_{n\in\mathcal{T}_{\mathrm{OOS}}^{(s)}} WP_n \log P_{niy_n}^{(s)}}{\sum_{n\in\mathcal{T}_{\mathrm{OOS}}^{(s)}} WP_n}. \tag{18}$$

- **Top-k accuracy (higher is better)** with $k = 10$, defined as the weighted share of observations for which the chosen destination is ranked among the k highest-probability destinations (Petersen et al., 2022).
- **Mean reciprocal rank (MRR; higher is better).** Let rank_n denote the rank position of the chosen destination in the predicted ordering. This metric is higher when the observed destination appears very high in the predicted ranking (Shi et al., 2012):

$$\mathrm{MRR} = \frac{\sum_{n\in\mathcal{T}_{\mathrm{OOS}}^{(s)}} WP_n \frac{1}{\mathrm{rank}_n}}{\sum_{n\in\mathcal{T}_{\mathrm{OOS}}^{(s)}} WP_n}. \tag{19}$$

- **Fitting factor (FF; higher is better).** As an additional descriptive probability-based summary, the fitting factor is defined as the weighted average predicted probability assigned to the observed destination (de Luca & Cantarella, 2016):

$$\mathrm{FF} = \frac{\sum_{n\in\mathcal{T}_{\mathrm{OOS}}^{(s)}} WP_n P_{niy_n}^{(s)}}{\sum_{n\in\mathcal{T}_{\mathrm{OOS}}^{(s)}} WP_n}. \tag{20}$$

- **McFadden's pseudo-R^2 (higher is better).** Defined as

$$R^2_{\mathrm{McF,OOS}} = 1 - \frac{\mathrm{LL}_{\mathrm{model}}}{\mathrm{LL}_0}, \tag{21}$$

  where $\mathrm{LL}_{\mathrm{model}}$ is the weighted OOS log-likelihood of the estimated model and LL_0 is the weighted log-likelihood of a null model with equal choice probabilities over the full destination set. This index measures how much the estimated model improves over an equal-probability null benchmark (Train, 2009).

Predicted destination shares are obtained by aggregating full-set predicted probabilities across all OOS observations. For each destination j, the model-implied weighted share is defined as

$$\widehat{S}_j = \frac{\sum_{n\in\mathcal{T}_{\mathrm{OOS}}^{(s)}} WP_n P_{nij}^{(s)}}{\sum_{n\in\mathcal{T}_{\mathrm{OOS}}^{(s)}} WP_n}, \tag{22}$$

and is compared with the corresponding observed weighted destination share S_j. The aggregate comparison includes:

- **Pearson and Spearman correlation (higher is better)** between S_j and $\widehat{S}_j$.

- **Jensen–Shannon (JS) divergence (lower is better)** between the observed and predicted destination-share distributions. This is a symmetric divergence measure between two probability distributions (Menéndez et al., 1997):

$$\mathrm{JS}(S, \widehat{S}) \;=\; \frac{1}{2}\sum_{j=1}^{J} S_j \log\left(\frac{S_j}{M_j}\right) \;+\; \frac{1}{2}\sum_{j=1}^{J} \widehat{S}_j \log\left(\frac{\widehat{S}_j}{M_j}\right), \tag{23}$$

  where M_j denotes the midpoint distribution, defined elementwise as

$$M_j \;=\; \frac{1}{2}\left(S_j + \widehat{S}_j\right). \tag{24}$$

Finally, I assess whether the model implies plausible travel-distance patterns. For each OOS observation, the full-set choice probabilities are used to compute the probability-weighted expected trip distance implied by the model, based on the zone-to-zone network-distance matrix. The resulting distribution of predicted trip distances is then compared with the observed OOS distance distribution. For long/day trips derived from *Tagesreisen*, the observed distance measure refers to the full reported trip and may therefore include both the outbound and return legs, whereas the model-implied distance is computed as a one-way origin–destination distance. As a result, comparisons of observed and predicted distance distributions for the long-trip segments should be interpreted with caution, since the observed distances may be mechanically larger.

The metrics described above are also used to compare alternative utility specifications under an identical experimental setting. In each benchmark comparison, the utility specification is re-estimated on the same training sample and evaluated on the same out-of-sample split, while the destination universe, observation weights, full-set probability computation, and all evaluation metrics are kept unchanged. The comparison is therefore directly attributable to differences in the utility formulation rather than to differences in the evaluation protocol.

1. **Proposed vs. accessibility-only baseline.** The proposed model is compared to an accessibility-only specification, $V_{ij} = \alpha\,\mathrm{EMU}_{ij}$, in order to quantify the incremental predictive contribution of the attractivity term.
2. **Proposed vs. operational SIMBA baseline.** The proposed specification is compared with the current operational SIMBA leisure destination utility, as described in Section 2.3.
3. **Proposed vs. pooled specifications.** The proposed segment-specific specification is compared with pooled variants based on the same destination-utility formulation, but estimated without the segmentation, in order to assess the predictive value of segmentation itself.

5.4 Interpretation of model coefficients

A further objective of this thesis is behavioural interpretation. The aim is to identify which destination characteristics are most strongly associated with leisure destination choice within each segment, and whether these patterns differ across segments. In the multinomial logit model, estimated coefficients enter the systematic utility function and should therefore be interpreted primarily as utility parameters. Their effects on choice probabilities are only indirect, since the corresponding probability changes depend on the full set of competing alternatives and their utilities. The coefficient tables reported in Chapter 6 present point estimates together with significance markers, which are used here as approximate indicators of coefficient precision. Approximate standard errors are obtained in a post-estimation step from a Hessian-based approximation of the weighted negative log-likelihood around the estimated optimum. Approximate z-statistics and two-sided p-values are derived under the asymptotic normal approximation (Peress, 2024). The corresponding inferential procedure, together with the construction of pairwise coefficient comparisons within and across segments, is described in Appendix I. Because all destination-side variables are transformed and standardised on a common scale, their coefficient magnitudes can be compared descriptively. By contrast, the accessibility coefficient is interpreted separately, since EMU_{ij} is a logsum-based accessibility measure but is not expressed on the same standardised scale. The interpretation of the estimated coefficients in this thesis is organised around four complementary aspects:

- **Sign.** The sign of a coefficient indicates whether a destination attribute is associated with higher or lower systematic utility, conditional on the other variables included in the specification. Positive coefficients imply that higher values of the attribute increase the systematic utility of a destination, while negative coefficients imply the opposite. Since several destination-side indicators are spatially correlated, signs are interpreted as conditional associations rather than as isolated causal effects (Train, 2009).
- **Magnitude.** Because all destination-side variables are transformed and then standardised as z-scores, their coefficients are expressed on a common scale. This makes coefficient magnitudes directly comparable across destination-side variables within the same segment. In particular, a one-unit increase in a standardised covariate corresponds to a one-standard-deviation increase (sd $= 1$) in the transformed destination attribute. In the utility function, this changes systematic utility by an amount equal to the corresponding coefficient. Larger absolute coefficients therefore indicate a stronger association with systematic utility and are especially informative for within-segment comparisons (Wooldridge, 2016).
- **Comparison across segments.** Since all destination-side variables are expressed on the same transformed and standardised scale in every segment-specific model,

coefficient magnitudes can also be compared descriptively across segments. These comparisons are used here to assess whether a given destination attribute appears more or less prominent for one segment than for another. At the same time, they should be interpreted with caution, because in logit-type models differences in coefficient magnitude may partly reflect differences in the scale of the unobserved component rather than only differences in underlying preferences (Williams, 2009).

- **Odds-based reading.** Because the model has a logit form, coefficients can also be read in terms of relative odds between two destination alternatives. Holding all other utility components constant, a one-standard-deviation increase in a destination attribute for one alternative multiplies the odds of choosing that destination rather than another by $\exp(\beta_h^{(s)})$. This provides an intuitive reading of coefficient magnitude. However, the implied change in actual choice probability still depends on the utilities of the full set of competing destinations (Train, 2009; Wooldridge, 2016).

5.5 Robustness analysis

The purpose of the robustness analysis is to assess the sensitivity of the estimated models to the main modelling choices taken. The aim is not to optimise the specification ex post, but to verify that the main conclusions do not depend on any single decision. To do so, I modify one component of the baseline specification at a time, while leaving all remaining elements unchanged, and re-estimate the model following the same procedure described in Section 5.3. The resulting alternative specifications are then compared to the baseline in terms of predictive performance under out-of-sample evaluation over the full destination universe. Differences in performance are assessed using the five core metrics introduced above, namely weighted NLL, MRR, McFadden's pseudo-R^2, Spearman correlation, and JS divergence. The robustness analysis is organised around the following families of checks:

- **R1: Robustness to preprocessing choices.** The first family of checks assesses whether the results depend on specific preprocessing decisions. I first test the definition of the destination universe by excluding Liechtenstein zones and all trips with origin or destination in Liechtenstein. I then estimate the model using raw destination-side indicators instead of the $\log(1+x)$-transformed variables. In a separate specification, I omit the global z-standardisation step. I also replace the baseline count-based specification with density-based versions of the four main amenity-count variables. In addition, I perform repeated leave-one-variable-out estimations, removing one of the 13 retained destination-side variables at a time. I further estimate a reduced specification in which the attractivity term is constructed using only the four most important destination-side variables. I then estimate the model after excluding observations classified in the *Visits* leisure category.

- **R2: Robustness to the accessibility component.** I estimate alternative models in which EMU_{ij} is replaced by simpler impedance proxies, namely the beeline distance between traffic-zone centroids and the network travel-distance skim.

- **R3: Robustness to sampled choice-set construction.** First, I vary the number of sampled non-chosen destinations for $K \in \{100, 300, 500, 800, 2000, 3000\}$. Second, I estimate an EMU-guided sampling strategy in which, for each trip, the non-chosen alternatives are drawn from a set of more plausible destinations on the basis of the accessibility distribution from the trip origin. The corresponding McFadden correction is applied in the sampled-choice likelihood.

- **R4: Robustness to the out-of-sample split design.** I estimate the model under an origin-based split, in which all trips originating from the same traffic zone are assigned jointly to the same side of the split.

- **R5: Robustness across survey years.** I estimate the model separately on the 2015 and 2021 survey waves.

- **R6: Robustness to regularisation.** Because the retained destination-side indicators are partly correlated, I additionally compare the baseline unpenalised model with two penalised variants: an L2-penalised (ridge) specification and an L1-penalised (lasso) specification (Friedman et al., 2010; Tibshirani, 1996). In both cases, the penalty is applied only to the destination-side coefficients β, while the accessibility coefficient is left unpenalised.

- **R7: Robustness to spatially lagged destination attributes and residual spatial autocorrelation.** I estimate an alternative specification that augments the baseline destination utility with a spatial spillover term constructed from spatially lagged destination attributes. More specifically, the utility is extended by adding a term of the form $(WX)_j\theta$, where WX denotes the spatial lag of the retained destination-side variables and W is a row-normalised contiguity matrix based on polygon intersection (Anselin, 2022). After full out-of-sample evaluation, I compute Moran's I for the residual destination shares, $r_j = S_j^{obs} - S_j^{pred}$, as a diagnostic of residual spatial autocorrelation. I also compare the dispersion of the spillover component $(WX)_j\theta$ with that of the direct destination component $X_j\beta$.

5.6 Integration and evaluation in SIMBA MOBi

A caveat is necessary regarding the SIMBA integration. The coefficients of the utility implemented in SIMBA MOBi were estimated from an earlier version of the short-trip leisure sample. Here, the return-home legs following leisure activities were inadvertently retained among the short leisure observations. This issue was identified only after the operational implementation had been completed, and it was not feasible within the thesis timeline to repeat the full integration using the corrected short-trip sample. For transparency, the coefficients actually used in SIMBA MOBi are reported in Appendix M. Although some coefficient differences are visible, the resulting attractivity indices and utility matrices remain highly correlated overall (Tables 32 and 33). The SIMBA exercise is therefore retained as a proof of concept and its interpretation should remain cautious, especially for the short-trip segments.

The estimated coefficients are used to construct a new segment-specific leisure destination utility for implementation in SIMBA MOBi in place of the current operational leisure destination-choice specification (see Section 2.3). The objective is to assess whether the offline-estimated utility can be translated into an operational specification and whether it yields plausible outcomes once embedded in the broader simulation framework. For each segment s, the estimated leisure destination utility is defined as

$$\widehat{V}_{ij}^{(s)} = \widehat{\alpha}^{(s)}\,\mathrm{EMU}_{ij} + \sum_{h=1}^{H} \widehat{\beta}_h^{(s)}\,\widetilde{F}_h(j), \tag{25}$$

where $\widehat{\alpha}^{(s)}$ and $\widehat{\beta}_h^{(s)}$ denote the estimated coefficients for segment s, EMU_{ij} is the accessibility component, and $\widetilde{F}_h(j)$ are the transformed and standardised destination-side variables.

For the integration experiment, only the leisure zonal destination-choice component is modified, while the remainder of the modelling system is kept unchanged. The proposed utility is thus evaluated without additional recalibration, in a setting where destination choice interacts with daily activity schedules, network conditions, and feasibility constraints.

Simulated agents are exposed to different leisure destination utilities depending on their segment. Age is taken directly from the agent attributes in the Synthetic Population and determines whether the young or old specification is applied. The trip-type dimension, by contrast, is not directly observable in the operational setting and is therefore approximated through the presence or absence of scheduled primary activities in the daily plan. In practice, agents without primary activities such as work or education are assigned to the long-trip utility, whereas the remaining agents are assigned to the short-trip utility. This short/long assignment should be interpreted as an operational proxy. Long leisure trips are more likely to occur on weekends, whereas the present SIMBA MOBi application simulates an average weekday. Moreover, exposure to the long-trip utility does not guarantee that a distant leisure destination will be realised, since realised choices remain constrained by the broader daily schedule, including time budgets and activity feasibility. The proxy is also not neutral across socio-demographic groups, as older agents are, by construction, more likely to have no scheduled primary activity and therefore more likely to be assigned to the long-trip utility.

On this basis, three simulation setups are considered: the current SIMBA MOBi reference leisure utility, applied with and without the operational correction terms (λ_j and λ_{ij}), and the proposed segmented leisure utility. In the reference specification, all agents are exposed to the same leisure destination utility, so the segment labels are used only for ex-post analysis of the realised trips. In the proposed specification, by contrast, agents are exposed to segment-specific leisure destination utilities, so the segmentation directly affects the simulated destination-choice process. The version without correction terms is retained as a secondary reference and is discussed only where it is useful for interpretation. For computational reasons, the proposed utility is applied only to a 10% subsample of the synthetic population, drawn uniformly at random. The simulation itself nevertheless continues to include the full set of agents. The empirical analysis is then based only on the leisure trips realised by this treated 10% subsample.

The outputs of interest are the leisure trips extracted after the simulation (MOBi.Sim). The evaluation logic remains intentionally simple and consistent with the offline benchmarking strategy. For each realised leisure trip in the analysed subsample, the retained information includes the origin zone, the destination zone, the travelled distance, and the segment to which the agent is assigned. These trip-level outputs are then aggregated and compared across simulation setups and against the Microcensus benchmark in terms of distance distributions and the spatial distribution of chosen destination shares.

6 Results

This chapter presents the empirical results of the thesis, following the methodological framework introduced in the previous chapter. Section 6.1 reports the results of the proposed destination-choice model, including the main evaluation metrics, the estimated destination shares, and their spatial patterns across the study area. It also presents and analyses the estimated coefficients. Section 6.2 compares the proposed specification against alternative model formulations. Section 6.3 summarises the results of the robustness analysis. Finally, Section 6.4 presents the results obtained from integrating the new utility specification into the operational SIMBA MOBi model.

6.1 Destination choice model results

This section presents the main empirical results of the estimated destination-choice models. It first reports their out-of-sample performance and the spatial patterns implied by the predicted destination shares. It then discusses the estimated coefficients of the accessibility component (EMU) and of the 13 destination-side attractivity variables. Additional diagnostics and supporting results are reported in Appendix J.

Model performance

Table 8 reports the main out-of-sample evaluation metrics under full choice-set evaluation. The upper panel summarises probability-based fit and ranking performance, while the lower panel reports aggregate destination-share diagnostics and distance-based plausibility indicators. These metrics are not straightforward to interpret in absolute terms, since their values depend on the size of the choice set and on the overall difficulty of the prediction task. In this study, they are therefore interpreted primarily in relative terms, by comparing results across segments and later across alternative model specifications.

Model performance differs across segments, with the most pronounced contrast emerging between short and long leisure trips. The two short-trip segments consistently achieve better results than the two long-trip segments across the reported metrics. Although part of this contrast is likely related to the much larger estimation samples available for short trips, it also suggests that long leisure trips are intrinsically more difficult to predict. The weakest performance is observed for the Old–Long segment, which is also the segment with the smallest number of observations. Across the probability-based fit and ranking metrics, the segment that shows the strongest overall performance is Old–Short, despite being estimated on a smaller sample than Young–Short. This may indicate that the behaviour captured in the Old–Short segment is more homogeneous and therefore easier to predict. By contrast, the Young–Short group likely captures a broader and more heterogeneous set of behaviours, given the wider age range it covers. By contrast, in

Table 8: Out-of-sample evaluation metrics for the proposed destination-choice model by segment under full choice-set evaluation.

(a) Fit and ranking metrics

Segment	NLL avg.	MRR	FF	McFadden R^2	Top10
YS	4.891	0.211	0.0539	0.456	0.447
OS	4.681	0.239	0.0701	0.479	0.490
YL	7.337	0.037	0.0028	0.183	0.092
OL	7.573	0.037	0.0018	0.157	0.081

(b) Aggregate and distance diagnostics

Segment	Pearson	Spearman	JS	$\bar{d}_{\text{obs}}$ [km]	$\bar{d}_{\text{pred}}$ [km]	$\Delta\bar{d}$ [km]
YS	0.622	0.488	0.163	12.79	12.84	0.05
OS	0.463	0.359	0.315	12.35	11.12	-1.23
YL	0.160	0.196	0.489	102.61	67.62	-34.99
OL	0.190	0.143	0.565	127.34	81.31	-46.03

the aggregate destination-share diagnostics, Young–Short emerges as the best-performing segment. This difference may itself partly reflect sample size, since aggregate metrics are likely to benefit from a larger empirical basis.

McFadden's R^2 shows that all four models improve over a random equal-probability benchmark, with the strongest gains observed for the short-trip segments. The ranking-based metrics point in the same direction. The *Top10* metric indicates that the chosen destination often appears near the top of the ranking of alternatives. For the two short-trip segments, around half of the observed destinations fall within the ten highest-probability destinations. The Pearson and Spearman correlations between observed and predicted destination shares indicate a reasonably strong spatial correspondence for the short-trip segments, whereas the relationship is clearly weaker for the long-trip segments. Even so, the correlation remains positive for both long-trip segments. The Jensen–Shannon divergence points in the same direction, although it is notably higher for Old–Short.

The comparison between observed and predicted trip distances should be interpreted differently across segments. For Young–Short, the model reproduces the average travelled distance very closely. For Old–Short, by contrast, the model tends to underestimate travelled distance, as the predicted mean remains below the observed one. For the two long-trip segments, the reported distance results are shown for completeness but are intrinsically difficult to interpret. In these segments, the observed distance measure derived from *Tagesreisen* may include both outbound and return legs, whereas the model-implied distance is computed as a one-way distance. As a result, the apparent underestimation for long trips cannot be interpreted straightforwardly. Since a direct comparison of average distances is therefore ambiguous for long trips, Appendix J.1 provides a more detailed analysis of the full observed and model-implied distance distributions.

Model-implied destination shares

Figure 5 maps the *model-implied destination shares* for the four behavioural segments across the full zone system. For each origin zone i, the estimated model yields a probability distribution over all destination zones j. The mapped share of each destination is then obtained by averaging these probabilities across all origins, thus assigning equal weight to every origin zone. The resulting values can be interpreted as percentages of destination choice under a *constant-production* assumption, and therefore sum to one within each segment across all destination zones. This differs from a real demand scenario, in which destination shares would also depend on the actual volume of trips produced by each origin. However, because NPVM traffic zones are intended to be broadly comparable units, this equal-origin weighting remains a useful descriptive device. Because the total probability mass is distributed across 7,978 zones, the mapped values are very small in absolute terms.

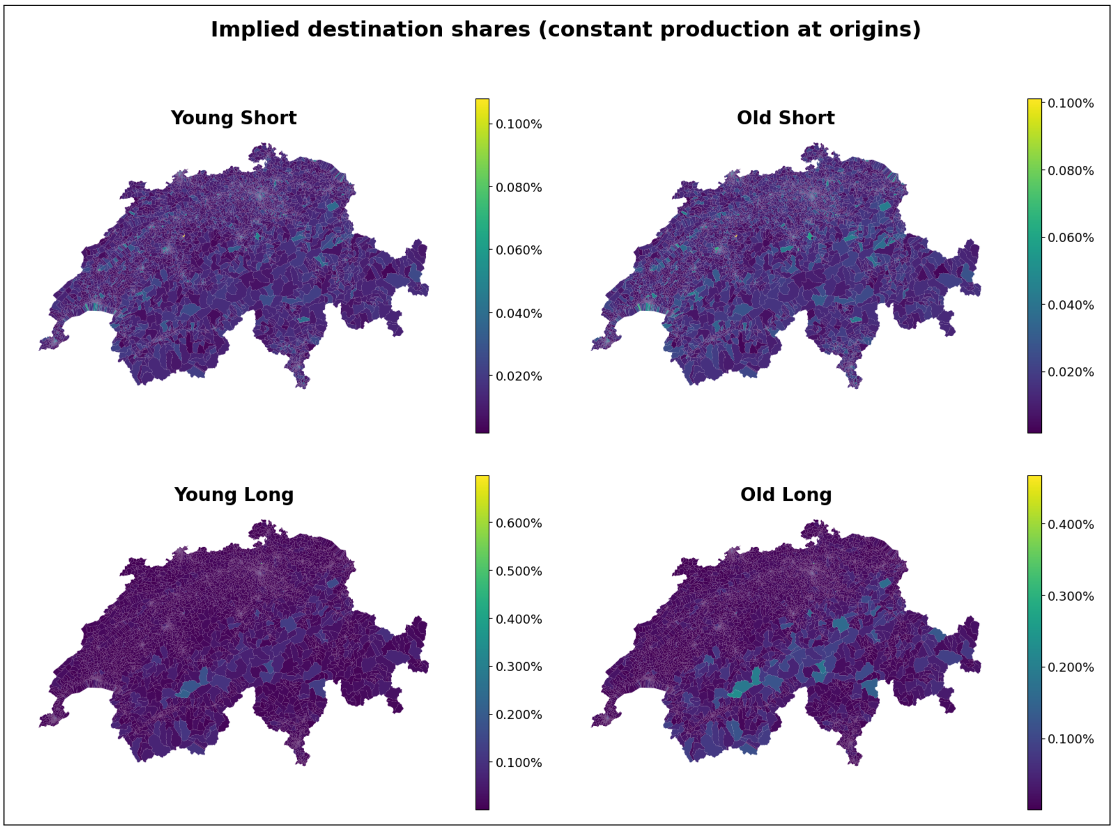


Figure 5: Model-implied destination shares by segment under constant origin production.

A striking pattern is the contrast between short and long trips. The two long-trip segments appear more polarised spatially, with probability mass concentrated in a more limited set of destinations. In both age groups, the largest implied destination shares for long trips are concentrated along the Alpine chain and in mountainous regions more generally. By contrast, the short-trip segments show a much more even distribution of implied destination shares. The two short-trip maps are relatively similar to one another, and age-related differences are not immediately pronounced at this aggregate national scale. Some local hotspots nevertheless emerge, especially near lakes and around several urban centres, including Bern and Lausanne. Overall, the maps suggest that short leisure mobility is associated with a broader and more spatially diffuse opportunity structure, whereas long leisure mobility is pulled more selectively towards a limited set of highly attractive destinations, especially in mountain areas. This is broadly in line with the patterns that had already emerged from the exploratory data analysis. A more detailed descriptive analysis is reported in Appendix J.2.

Estimated coefficients

The estimated coefficients of the final destination-choice specification are reported in Table 9. For compactness, the table presents only the coefficient values together with approximate significance markers (statistically distinguishable from zero). The accessibility coefficient α_{EMU} is shown separately at the bottom of the table. Additional coefficient details, including standard errors, z-statistics, and p-values, are reported in Appendix J.3. The appendix also provides further elements supporting coefficient interpretation, reporting the correlation structure among the retained destination-side features and their correlation with a destination-level accessibility summary derived from the EMU matrix. The interpretation that follows focuses primarily on coefficients that are statistically distinguishable from zero.

Table 9: Estimated model coefficients by segment. Significance markers denote coefficients approximately distinguishable from zero: *** for $p < 0.001$, ** for $p < 0.01$, and * for $p < 0.05$.

Variable	**YS**	**OS**	**YL**	**OL**
1. Gastronomy POIs	0.327***	0.347***	0.612***	0.567***
2. Resident population	-0.056***	-0.015	-0.061***	-0.037
3. Lake shore density	0.060***	0.079***	0.092***	0.112***
4. Hard outdoor POIs	-0.010	-0.051**	0.074***	0.046
5. Soft outdoor POIs	-0.048***	-0.016	-0.194***	-0.166***
6. Land-use mix	0.055***	0.040**	-0.152***	-0.130***
7. Cultural POIs	0.023***	0.037**	-0.094***	-0.102***
8. Sport POIs	0.112***	0.053**	-0.060**	-0.155***
9. Outdoor sport route length	0.200***	0.278***	0.480***	0.533***
10. Other leisure POIs	-0.005	-0.001	0.090***	0.045
11. Urban diversity index	0.051***	0.039	0.178***	0.289***
12. Urban POIs density	0.039**	-0.084**	0.037	0.036
13. Support/service amenities	0.181***	0.220***	0.265***	0.251***
Accessibility α_{EMU}	1.294***	1.345***	0.724***	0.634***

A first general result concerns the accessibility coefficient α_{EMU}, which is positive and statistically distinguishable from zero in all four segment-specific models. Although its magnitude is not directly comparable to that of the destination-side coefficients, it is nonetheless large on its own scale. This indicates that accessibility constitutes a major component of the systematic utility in all segments. Its role is clearly stronger for short trips than for long trips, with estimated coefficients that are substantially larger in the two short-trip segments. Across age groups, the coefficient is slightly smaller for younger individuals in the short-trip segment. For long trips, the pattern is reversed. These age differences remain modest in magnitude.

Sign Analysis A first general result is that coefficient signs remain broadly stable across age groups within the same trip-type category. By contrast, the main sign reversals emerge between short- and long-trip segments. Some features also display a consistently positive sign across all segments, most notably *Gastronomy POIs (1)*, *Lake shore density (3)*, *Outdoor sport route length (9)*, and *Support/service amenities (13)*. Their stable sign suggests these are consistently associated with higher destination utility for leisure travel.

For short trips, the sign pattern is predominantly positive. The three negative exceptions are *Resident population (2)* and *Soft outdoor POIs (5)* for the Young–Short segment, and *Hard outdoor POIs (4)* for the Old–Short segment. A more notable contrast concerns *Urban POIs density (12)*, which shows a significant sign reversal across the two short-trip age groups: it is positive in Young–Short and negative in Old–Short. This suggests that, for younger individuals, zones with a high density of urban POIs are associated with higher utility, whereas for older individuals the association goes in the opposite direction. At the same time, this variable is also the one most strongly correlated with accessibility.

For long trips, the sign structure is more mixed, with several destination-side variables showing negative coefficients. These include *Land-use mix (6)*, *Cultural POIs (7)*, *Sport POIs (8)*, and *Soft outdoor POIs (5)*. Taken together, these results are broadly consistent with the spatial pattern of long-trip destination shares, which are concentrated in Alpine and mountain zones where these features are typically less prevalent than in more urban or peri-urban areas. A more nuanced case is *Land-use mix (6)*. Although a negative sign may initially seem counterintuitive, in Alpine areas the limited built component tends to reduce this index. By contrast, *Urban diversity index (11)* is positive and significant in both long-trip segments. This suggests that, for long trips, a broad and diversified urban-type offer may still contribute positively to utility. This interpretation should remain cautious because the variable is relatively composite and also has the highest VIF.

Magnitude analysis The analysis of coefficient magnitudes helps distinguish the more central features of destination utility from more marginal ones within each segment-specific model. Several general patterns should be kept in mind. First, only a subset of the retained destination-side variables shows coefficient magnitudes large enough to suggest a substantively important role in destination utility. Second, relatively small coefficients do not necessarily imply that the corresponding destination feature is unimportant in a broader sense. Part of its explanatory content may already be absorbed by the accessibility term or by other correlated destination-side variables. Third, destination-side coefficients tend to be descriptively larger in absolute value in the long-trip models than in the short-trip models. A further notable point is that *Resident population (2)* remains comparatively small in magnitude and, for the older groups, is not statistically distinguishable from zero. This is not entirely surprising, given that population is distributed rather broadly across the traffic-zone system.

For the Young–Short segment, the most notable coefficients are *Gastronomy POIs (1)*, *Support/service amenities (13)*, and *Outdoor sport route length (9)*, followed by *Sport POIs (8)*. Among these, *Gastronomy POIs (1)* has the largest positive coefficient. Its correlation with destination-side accessibility summary is weak, which indicates that its role is not simply capturing highly accessible destinations. At the same time, it is moderately correlated with *Cultural POIs (7)* and *Urban diversity index (11)*, suggesting that it may also reflect a broader destination profile. The second-largest coefficient is *Support/service amenities (13)*, followed closely by *Outdoor sport route length (9)*, whose difference in magnitude is not statistically significant. *Outdoor sport route length (9)* is strongly negatively correlated with destination-side accessibility, suggesting that it partly captures the appeal of less accessible destinations.

For the Old–Short segment, a similar general group of variables remains central, but their relative importance changes. The largest coefficient is *Gastronomy POIs (1)*, followed by *Outdoor sport route length (9)* and *Support/service amenities (13)*. Among these three coefficients, the only statistically significant pairwise difference is the one between *Gastronomy POIs (1)* and *Support/service amenities (13)*. This suggests that, overall, these three features contribute in a relatively balanced way to destination utility in this segment.

The Young–Long segment exhibits a more polarised coefficient structure. The most notable positive coefficients are *Gastronomy POIs (1)*, *Outdoor sport route length (9)*, *Support/service amenities (13)*, and *Urban diversity index (11)*, while the most notable negative ones are *Soft outdoor POIs (5)* and *Land-use mix (6)*. Several of these variables also appeared in the short-trip models, but here they reach significantly larger magnitudes than in the short-trip models. The largest coefficient is again attached to *Gastronomy POIs (1)*. Although gastronomy is not a dominant stated purpose among long leisure trips, this may indicate that gastronomic provision matters as a complementary attribute of broader destinations rather than as the primary trip motive. The strong positive coefficient for *Outdoor sport route length (9)* is more directly intuitive, as route-based outdoor infrastructure is closely associated with excursion-type leisure. *Support/service amenities (13)* also remains important, suggesting that many long-trip destinations require a minimum level of supporting services. On the negative side, the magnitudes of *Soft outdoor*

POIs (5) and *Land-use mix (6)* suggest that long-trip attractivity is shaped not only by the presence of appealing features, but also by the relative absence of more everyday environments.

For the Old–Long segment, the overall pattern is very similar. The strongest positive coefficients are again *Gastronomy POIs (1)*, *Outdoor sport route length (9)*, *Urban diversity index (11)*, and *Support/service amenities (13)*, while *Soft outdoor POIs (5)*, *Sport POIs (8)*, *Land-use mix (6)*, and *Cultural POIs (7)* show the largest negative magnitudes. The resulting destination profile is therefore close to that of Young–Long, with only moderate differences in ordering and intensity. One notable feature is the larger role of *Lake shore density (3)*, which reaches a higher positive magnitude in this segment than in the others. Another interesting aspect is the coexistence of a strongly positive coefficient for *Urban diversity index (11)* and negative coefficients for *Cultural POIs (7)* and *Sport POIs (8)*, which suggests that what matters may be less the isolated abundance of specific amenities than the presence of a more diversified service bundle. More generally, the long-trip segments confirm that destination-side coefficients tend to have larger magnitudes.

Comparison across segments The comparison across segments confirms that the main source of heterogeneity is the contrast between short and long leisure trips rather than the contrast between age groups. Several coefficients change substantially when moving from the short-trip to the long-trip models within both age classes. By contrast, age-group differences are more limited and less systematic when the segments are compared directly. Even so, the overall coefficient configuration still suggests some meaningful differences between younger and older individuals within trip type.

- ***Sport POIs (8).*** This variable shows one of the clearest age-related contrasts. In the short-trip segments, the coefficient is positive for both age groups but larger for *Young–Short* than for *Old–Short*. In the long-trip segments, it becomes negative for both groups, with a stronger negative magnitude for *Old–Long*.
- ***Outdoor sport route length (9).*** Route-based sport and outdoor infrastructure is positive and important in all four segments. Within the short-trip models, the coefficient is larger for *Old–Short* than for *Young–Short*, suggesting a somewhat stronger role in the older group.
- ***Urban POIs density (12).*** For short trips, this variable displays a sign reversal between age groups: it is positive in *Young–Short* and negative in *Old–Short*, and in both cases the coefficient is significantly different from zero. This makes this variable one of the clearest age-differentiating variables in the short-trip models.
- ***Support/service amenities (13).*** Support infrastructure is positive in all four segments, but within the short-trip models the coefficient is larger for *Old–Short* than for *Young–Short*. This suggests a slightly stronger estimated association for

older individuals in short leisure travel.

Odds-ratio illustration As a final complementary step, a small set of coefficients is illustrated in odds-ratio terms in order to provide a more intuitive sense of magnitude. I focus on two destination-side variables that are especially relevant in the previous discussion and report, for each of them, the odds multiplier associated with a one-standard-deviation increase in the transformed covariate. For the two variables considered here, which enter the model as $z(\log(1 + x))$, I also provide an approximate translation back to the raw scale, using the median of the empirical distribution as the reference point.

- ***Sport POIs (8).*** Sport amenities show a clear contrast between short- and long-trip segments. A one-standard-deviation increase corresponds to an odds increase of about 11.9% for *Young–Short* and 5.4% for *Old–Short*, but to an odds reduction of about 5.8% for *Young–Long* and 14.4% for *Old–Long*. At the median of the raw distribution, this +1 standard-deviation change corresponds approximately to moving from 3.0 to 9.26 sport amenities in a zone, i.e. a difference of about 6.26 units.
- ***Support/service amenities (13).*** Support infrastructure is positive in all four segments, but its odds effect is somewhat stronger in the long-trip models. A one-standard-deviation increase corresponds to an odds increase of about 19.8% for *Young–Short*, 24.6% for *Old–Short*, 30.3% for *Young–Long*, and 28.5% for *Old–Long*. At the median of the raw distribution, this change corresponds approximately to moving from 3.0 to 8.38 support amenities in a zone, i.e. a difference of about 5.38 units.

6.2 Benchmark comparison with alternative specifications

To benchmark the proposed destination-choice specification, I compare it with three alternative references estimated on the same training and out-of-sample splits, but with different utility specifications: an *Accessibility-only* model, the current *SIMBA baseline* specification, and pooled versions of the proposed model estimated without the full segment-specific differentiation. Table 10 reports the main fit, ranking, aggregate-share, and distance-plausibility metrics. The first four panels compare the *Proposed*, *SIMBA*, and *Accessibility-only* specifications separately for each segment. Within each segment-specific panel, the best value across the three specifications is highlighted in bold. The final panel reports the results of the pooled models estimated without age segmentation: *Short-only*, *Long-only*, and *All* (Short + Long) models. The benchmark comparison shows that the proposed destination-attractivity specification is competitive with, and often superior to, the current SIMBA MOBi specification.

For the two short-trip segments, the probability-fit measures improve only modestly. The improvement is slightly more visible in the aggregate-share diagnostics. In particular,

Table 10: Benchmark comparison across model specifications under full out-of-sample evaluation.

YS (Young, Short)

Model	NLL	MRR	R^2	Top10	Pearson	Spearman	JS	$\Delta\bar{d}$ [km]
Proposed	**4.891**	**0.211**	**0.456**	**0.447**	**0.622**	**0.488**	**0.163**	0.048
SIMBA	4.943	0.209	0.450	0.446	0.560	0.455	0.174	0.367
Acc.-only	5.036	0.208	0.439	0.430	0.437	0.377	0.198	**-0.002**

OS (Old, Short)

Model	NLL	MRR	R^2	Top10	Pearson	Spearman	JS	$\Delta\bar{d}$ [km]
Proposed	**4.681**	**0.239**	**0.479**	**0.490**	**0.463**	**0.359**	**0.315**	-1.227
SIMBA	4.763	0.227	0.470	0.482	0.417	0.328	0.332	**-0.669**
Acc.-only	4.823	0.232	0.463	0.470	0.377	0.297	0.349	-0.807

YL (Young, Long)

Model	NLL	MRR	R^2	Top10	Pearson	Spearman	JS	$\Delta\bar{d}$ [km]
Proposed	**7.337**	**0.037**	**0.183**	**0.092**	**0.160**	**0.196**	**0.489**	-34.995
SIMBA	7.618	0.032	0.152	0.065	0.141	0.048	0.540	-34.852
Acc.-only	7.940	0.031	0.116	0.063	0.045	-0.072	0.557	-37.520

OL (Old, Long)

Model	NLL	MRR	R^2	Top10	Pearson	Spearman	JS	$\Delta\bar{d}$ [km]
Proposed	**7.573**	**0.037**	**0.157**	**0.081**	**0.190**	**0.143**	**0.565**	-46.033
SIMBA	8.060	0.017	0.103	0.041	0.091	0.041	0.614	-47.694
Acc.-only	8.298	0.020	0.076	0.039	0.008	-0.042	0.627	-49.377

Pooled results (age aggregation)

Group	NLL	MRR	R^2	Top10	Pearson	Spearman	JS	$\Delta\bar{d}$ [km]
Short-only	4.854	0.217	0.460	0.455	0.644	0.511	0.138	-0.243
Long-only	7.379	0.036	0.179	0.086	0.169	0.206	0.471	-36.458
All	5.360	0.185	0.403	0.395	0.510	0.487	0.142	-7.606

the proposed specification yields higher Pearson and Spearman correlations and lower Jensen–Shannon divergence. A similar pattern emerges when comparing the proposed specification with the accessibility-only benchmark: even for short leisure trips, which are already strongly explained by accessibility alone, introducing destination-side attractivity improves overall performance only moderately. At the same time, distance plausibility remains slightly better captured by the benchmark models. For the Young–Short segment, the accessibility-only model comes closest to the observed mean distance, although the difference from the proposed specification is very small. For the Old–Short segment, the

SIMBA baseline performs better and the gap relative to the proposed model is somewhat larger. Taken together, these comparisons suggest that, within the short-trip segments, the added value of the proposed specification is present but limited, and appears somewhat more pronounced for older individuals than for younger ones.

For the long-trip segments, the advantage of the proposed specification is much clearer. Although absolute performance remains lower than for short trips, the gains over both accessibility-only and SIMBA are larger. Here, the proposed model improves all core metrics. The gains are especially pronounced in the share-based metrics, indicating that the proposed attractivity structure captures the geography of long-distance leisure destinations more successfully than the simpler alternatives. This is particularly important because the accessibility-only benchmark performs poorly in these segments. In this sense, the benchmark results suggest that the proposed specification adds its greatest value precisely where destination choice is less directly constrained by accessibility.

A potential limitation of the proposed specification is that it is more complex than the SIMBA baseline, since it includes thirteen destination-side variables rather than the three high-level components currently implemented in SIMBA MOBi. However, an additional reduced-form experiment shows that a substantially more compact specification can still retain a meaningful share of the predictive performance of the full model (Section 6.3). A model using only four destination-side indicators performs slightly worse than the full specification, but remains reasonably competitive on the main fit and share-reproduction metrics across segments. In principle, a more parsimonious operational version could therefore be constructed if simplicity and ease of implementation were prioritised more strongly than maximising predictive performance.

Finally, the pooled benchmarking results provide further support for the usefulness of segmentation. When the model is estimated on pooled *Short-only*, *Long-only*, or fully aggregated *All* models, the results suggest that segmentation helps capture heterogeneity in leisure destination choice. The pooled *Short-only* model performs slightly better than the Young–Short model but remains weaker than the Old–Short model. By contrast, this pattern is much less visible for long trips, where pooling leads only to very limited differences. When all leisure trips are estimated together in a fully aggregated *All* model, performance deteriorates relative to the short-segmented specifications, although this comparison is mainly illustrative. Overall, these results support the view that segmentation is useful because it allows the model to adapt to systematically different destination-choice structures. For trip type, this conclusion is unsurprising, since short and long leisure trips appear to correspond to clearly different behavioural contexts, as also reflected in the coefficient patterns. For age, the evidence is more moderate, but still suggests that some behavioural heterogeneity is captured when the model is estimated separately by segment.

A further consideration concerns the age segmentation itself. Some residual heterogeneity within the younger group is plausible. For this reason, I also report a sensitivity analysis based on a finer age segmentation for short trips, where the sample size is sufficient to support a more granular split (Appendix K).

6.3 Robustness analysis results

This section reports the main findings of the robustness analysis introduced in Section 5.5. Following the same structure introduced there, the discussion is organised into the seven robustness families $R1$–$R7$. Overall, the baseline specification proves to be generally robust: most alternative modelling choices produce only limited changes in out-of-sample performance. The few more visible deviations, mainly related to the $\log(1 + x)$ transformation and the exclusion of *Visits*, remain limited and do not overturn the main empirical conclusions. To keep the discussion concise, the main text focuses only on the substantive takeaways, while the complete metric tables are reported in Appendix L. Table 11 summarises the main findings.

Table 11: Summary of the main findings from the robustness analysis.

Family	Alternative specification tested	Main finding	Overall
R1	Preprocessing choices	The clearest deviations arise when dropping the $\log(1 + x)$ transformation or excluding *Visits*.	Mostly stable
R2	Alternative accessibility proxies	Simpler distance-based accessibility proxies yield broadly similar results.	Stable
R3	Sampled choice-set construction	Performance is stable across alternative K values and using EMU-weighted sampling.	Stable
R4	Out-of-sample split design	Deviations are somewhat larger for OL.	Stable
R5	Survey-year sensitivity	Only limited changes relative to the pooled baseline.	Stable
R6	Regularisation	L_1 and L_2 penalties have virtually no effect on predictive performance.	Negligible effect
R7	Spatial spillovers and residual dependence	Spatial spillovers add limited explanatory value and do not materially reduce residual spatial autocorrelation.	Limited additional gain

R1: Preprocessing choices. Overall, the baseline specification proves highly robust to most preprocessing choices. In most cases, the differences in performance are very small. Two preprocessing choices appear more consequential than the others. First, replacing the $\log(1+x)$ transformation with raw destination-side variables leads to a clearer deterioration in performance, especially for long trips. This suggests that the logarithmic transformation helps stabilise the model. Second, excluding *Visits* observations produces the most visible changes within this robustness family. In particular, removing *Visits* generally improves the probability-based fit measures, which is consistent with the idea that socially motivated trips are only imperfectly captured by a destination-choice model based on territorial attributes alone. At the same time, the coefficient patterns remain broadly stable under this robustness check, as shown in Table 30. The most noteworthy pattern is the further decline of the coefficient associated with *Resident population (2)*, which becomes more negative in all four segments once *Visits* are excluded. This is consistent with the interpretation that, in the baseline model, population partly helps capture socially motivated trips.

R2: Accessibility component. The baseline results remain broadly stable when EMU_{ij} is replaced by simpler impedance measures. Using the logarithm of travel-distance skims leads to only limited changes in performance, especially for the short-trip segments. Replacing EMU with log-beeline distance yields somewhat larger deviations across segments. Even so, the overall changes remain moderate, and the qualitative interpretation of the model is preserved.

R3: Sampled choice-set construction. Varying K over a wide range does not materially affect full-choice-set out-of-sample performance, indicating that the substantive conclusions do not depend on the exact size of the sampled choice set. Replacing uniform sampling with EMU-weighted sampling does not yield meaningful improvements either. Figure 6 provides an additional visual check based on negative log-likelihood. As K increases, sampled train and sampled OOS NLL move upward toward the full-set OOS level while remaining very close to each other within each segment, suggesting no evidence of meaningful overfitting.

R4: Out-of-sample split design. The alternative out-of-sample design based on a by-origin split produces limited changes relative to the baseline by-person split for most segments. The clearest deviation appears for the Old–Long segment, where predictive performance deteriorates more visibly, while the Old–Short segment shows some improvement. Overall, this robustness check does not alter the main conclusions.

R5: Survey year. Estimating the model separately on the 2015 and 2021 waves leads only to limited changes relative to the pooled baseline. Some differences across years are visible, but they remain moderate and do not alter the main empirical conclusions. Overall, the results suggest that the baseline specification is reasonably stable over time, despite some temporal heterogeneity between the two survey waves.

R6: Regularisation. The regularisation checks indicate only small changes in predictive performance when mild L_1 or L_2 penalties are applied to the destination-side

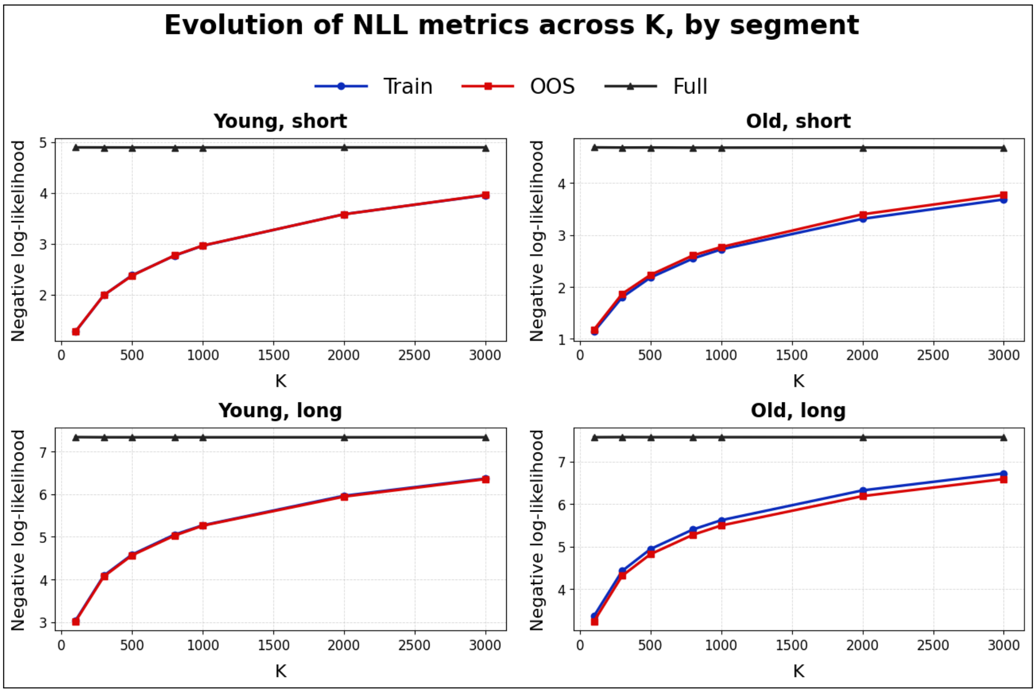


Figure 6: Evolution of training, sampled OOS, and full-choice-set negative log-likelihood across alternative values of K, by segment.

coefficients. The baseline specification therefore does not appear to be driven by unstable destination-side parameters.

R7: Spatial spillovers and residual spatial dependence. Allowing for spatial spillovers leads only to limited changes relative to the baseline specification. The spillover model yields some improvement in predictive performance, especially for the long-trip segments. Residual spatial autocorrelation is already weak in the baseline model, with Moran's I close to zero in all segments. After introducing the spillover term, the changes in Moran's I remain very small and are not systematic: it declines in YS (from 0.0070 to 0.0021) and OS (from -0.0059 to -0.0070), but moves slightly towards zero in YL (from -0.0158 to -0.0146) and OL (from -0.0094 to -0.0085). In addition, the spillover component $(WX)_j\theta$ remains less dispersed across destinations than the direct destination component $X_j\beta$, suggesting that the additional spatially lagged term captures only a more limited share of the overall variation in destination utility. Overall, the spillover specification provides limited additional predictive value and does not materially reduce residual spatial dependence. It therefore does not challenge the baseline model.

6.4 Integration in SIMBA MOBi and simulation results

This section reports the results of the proof-of-concept integration of the proposed segmented leisure utility into SIMBA MOBi. The analysis is based on the realised leisure trips extracted from the simulation outputs (MOBi.Sim) for the 10% subsample of treated agents. After assigning each simulated agent to one of the four segments (Young–Short, Old–Short, Young–Long, Old–Long), the extraction yields the segment-specific numbers of leisure trips reported in Table 12. These counts provide the basis for the subsequent comparison of distance distributions and spatial destination patterns across simulation setups and against the Microcensus benchmark.

Table 12: Number of extracted realised leisure trips by segment and simulation setup.

Segment	Reference	Reference w/o corr. terms	Proposed
Young–Short	349,662	349,858	346,882
Old–Short	7,240	6,915	6,927
Young–Long	137,006	136,137	135,115
Old–Long	124,122	124,249	123,903

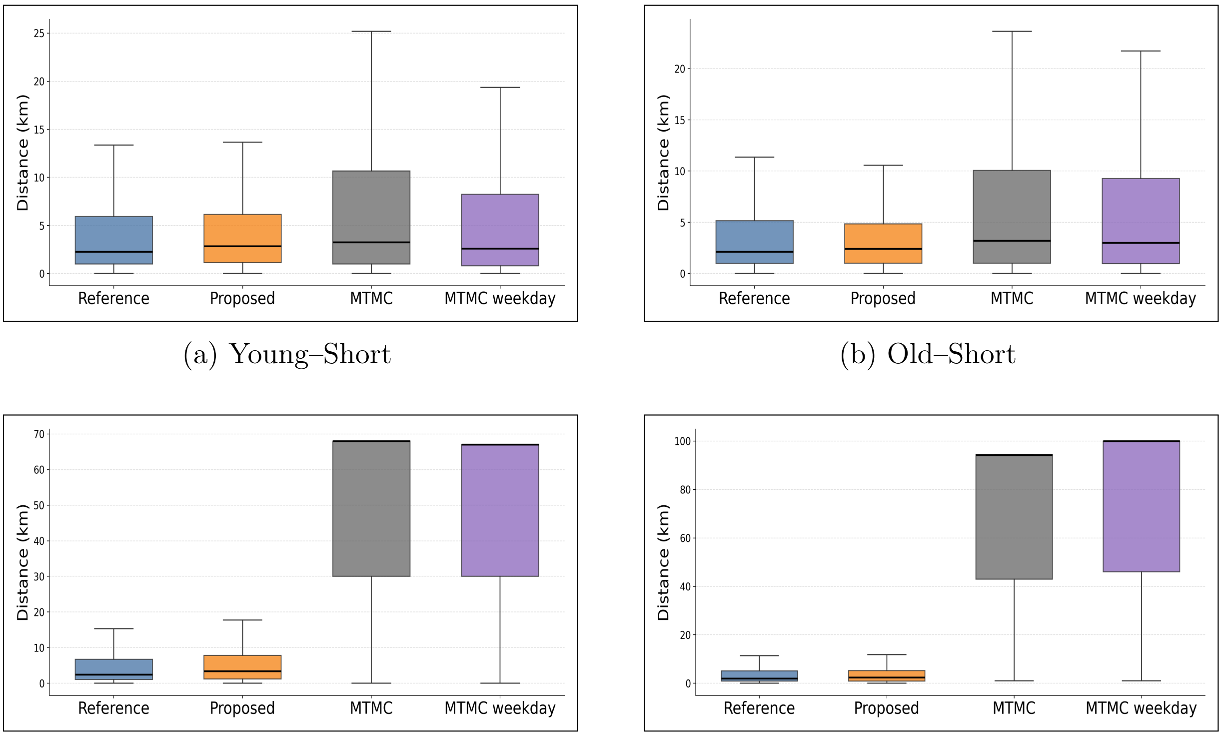
The extracted numbers of leisure trips are very similar across the available simulation setups. At the same time, the segment-specific trip counts are highly uneven. The Young–Short segment clearly contains the largest number of observations, whereas Old–Short is by far the smallest. This imbalance is consistent with the operational proxy used for the short/long assignment. In particular, older agents are much less likely to be assigned to the short-trip specification. A similar issue, though less extreme, also contributes to the comparatively large size of the two long-trip segments. These differences do not invalidate the simulation-based comparison, but they do imply that the Old–Short segment should be interpreted with greater caution. More generally, they confirm that the operational translation of the empirical segmentation remains one of the main limitations of the SIMBA integration experiment.

A further issue emerged in the spatial analysis of simulated destination shares using the proposed utility. In the simulated destination-share distributions, a very small number of zones (roughly three to four) display consistently and disproportionately high observed shares across multiple segments. These zones do not correspond to destinations that stand out in the offline model-implied share distributions and therefore cannot be straightforwardly interpreted as substantively plausible model outcomes. At present, their origin remains uncertain. They may reflect technical artefacts arising somewhere along the operational pipeline, for example in the assignment of agents to zones, in intermediate encodings, or in other processing steps. Importantly, these outlying zones are retained in the quantitative analysis, so the reported comparisons are based on the raw simulation outputs as produced. However, for cartographic visualisation only, their displayed values are capped in order to improve the readability of the maps.

The analysis is structured as follows. It first examines the distribution of simulated leisure-trip distances. It then analyses the zonal destination-share distributions associated with the chosen destinations. Supporting maps for this second part are provided in Appendix N. In all cases, the analysis is reported for the four segments.

Analysis of distance distributions

For the analysis of distance distributions, I do not discuss the reference setup without correction terms. Figure 7 summarises the distribution of realised leisure-trip distances across the different segment groups, while Table 13 reports the corresponding summary statistics. In each case, the simulated outputs are compared with both the full Microcensus benchmark and the weekday-only Microcensus benchmark, the latter being included to improve consistency with the weekday-based SIMBA MOBi simulation setting. The distance-distribution analysis shows that both simulation setups underestimate realised leisure-trip distances relative to the MTMC benchmark. The magnitude of this discrepancy varies across segments and is particularly pronounced for the long-trip groups.



(a) Young–Short (b) Old–Short

(c) Young–Long (d) Old–Long

Figure 7: Boxplots of realised leisure-trip distance distributions across simulation outputs and Microcensus benchmarks for the four operational segments. For visual readability, the display range is truncated in the long-trip panels (Young–Long and Old–Long) for the Microcensus and Microcensus weekday benchmarks.

The clearest result concerns the distinction between short and long trips. In the MTMC, the long-trip segments (*YL* and *OL*) are characterised by much larger mean, median, and upper-tail distances than the short-trip segments. It should be noted, however, that the MTMC distance measure refers to round trips, whereas the simulated distances are based on realised one-way leisure trips. Even with this difference in definition, the underestimation in the simulation outputs remains evident. The simulated distances for the long-trip segments remain close to those of the short-trip segments, suggesting that the

Table 13: Summary of realised leisure-trip distance distributions by segment. Reported values are the mean, median, and 90th percentile of travelled distance (km).

Source	**YS**	**OS**	**YL**	**OL**
Mean				
Reference	7.12	6.45	8.14	7.04
Proposed	7.99	5.56	9.68	7.22
MTMC	12.02	11.86	100.78	125.19
MTMC weekday	9.73	11.29	101.63	130.09
Median				
Reference	2.26	2.13	2.47	2.02
Proposed	2.85	2.42	3.39	2.42
MTMC	3.26	3.20	68.00	94.50
MTMC weekday	2.60	3.00	67.00	100.00
90th percentile				
Reference	15.07	11.79	18.78	12.01
Proposed	11.82	8.39	18.01	9.34
MTMC	30.70	28.53	240.00	273.00
MTMC weekday	23.73	26.34	244.00	292.80

operational implementation of the long-trip utility within SIMBA MOBi does not fully reproduce the strong separation observed in the survey data. A plausible explanation is that the long-trip utility operates within a simulation environment that still imposes substantial daily scheduling and feasibility constraints, limiting the extent to which long leisure excursions can emerge in practice.

The short-trip segments also remain too short relative to the MTMC benchmark, although the discrepancy is much smaller in magnitude. For *YS* and *OS*, the simulated mean and median distances are not dramatically far from the observed values, and for *YS* the median is even slightly higher. The upper tail of the distribution remains insufficiently represented, however. The simulated 90th percentiles for *YS* and *OS* remain well below the corresponding MTMC values. The segmented utility can thus influence realised trip-length distributions within SIMBA MOBi even without additional calibration. At the same time, the improvement remains limited. For *OS*, the *Proposed* specification yields a lower mean distance than the Reference. Moreover, with respect to the upper tail, the Reference specification remains closer to the MTMC values.

Overall, the distance-distribution analysis indicates that the *Proposed* utility alone is not sufficient to realign simulated leisure-trip distances with the MTMC benchmark, especially for long trips. At the same time, the results suggest that the integration does not disrupt the simulated distance structure and that, for some cases such as *YS*, the changes move in a plausible direction. The simulation results should therefore be interpreted not as a final calibrated improvement, but as proof-of-concept evidence that the proposed segmented destination utility can be integrated into the operational SIMBA environment. The effect of the segmentation is visible in the fact that the differences in travelled distances between young and old individuals are larger under the *Proposed* specification in both trip-type groups.

Analysis of destination shares

To analyse the spatial distribution of simulated leisure destinations, I compare the zonal destination-share distributions of the different specifications against the MTMC benchmark using Spearman correlation. Table 14 reports these correlations for four cases: the operational *Reference* specification, the *Reference* specification without correction terms, the *Proposed* specification integrated into SIMBA MOBi, and the *Proposed Offline* specification. The latter corresponds to the destination-share distributions implied directly by the estimated utility outside the SIMBA environment and is included only as a reference point. This comparison therefore helps distinguish two aspects: first, how closely the simulated destination shares match the observed MTMC spatial pattern; and second, how much of the offline spatial fit is preserved once the estimated destination utility is embedded in the broader SIMBA MOBi system. For a visual inspection of the corresponding spatial distributions, see Appendix N.

Table 14: Spearman correlations with the MTMC benchmark by segment.

Segment	Reference	Reference w/o corr.	Proposed	Proposed Offline
YS	0.5858	0.5854	0.3747	0.6443
OS	0.2298	0.2217	0.2027	0.4523
YL	0.0036	0.0039	0.2589	0.2510
OL	0.0770	0.0678	0.2317	0.2259

Overall, the *Reference* specification performs better in the short-trip segments, most clearly in *YS*, where it remains substantially closer to the MTMC benchmark than the integrated *Proposed* setup. By contrast, the *Proposed* specification performs better in the two long-trip segments, *Young–Long* and *Old–Long*. For *Old–Short*, the difference between the two integrated specifications is very small and, given the limited number of simulated trips in this group, should not be over-interpreted. Taken together, the results do not indicate a general improvement in spatial fit under the integrated *Proposed* specification, but they do point to some relative gains for the long-trip segments.

A useful reference point is provided by the offline model-implied destination shares. In the offline comparison, the *Proposed* specification performs better in all four segments. This contrast suggests that good offline performance does not automatically carry over to equally strong performance once the same utility is embedded in the full SIMBA MOBi environment. At the same time, this should not be interpreted as evidence that the proposed attractivity formulation is intrinsically weaker. Rather, it indicates that the estimated destination utility, in its current form, is not yet fully adapted to the operational simulation context.

The comparison between the *Reference* setup with and without correction terms further suggests that, in this benchmark, the correction terms alone are unlikely to explain most of the difference, since their correlations remain very similar across all four segments. More importantly, however, the operational *Reference* specification is not defined only by those correction terms. Its broader utility structure, including coefficients and scaling choices, has already been adapted to the SIMBA MOBi context. This likely gives the reference setup an important advantage in reproducing observed destination shares. For this reason, the weaker performance of the integrated *Proposed* specification is better understood as a consequence of inserting the proposed utility without an equivalent context-specific recalibration of the surrounding utility specification.

The most encouraging result concerns the two long-trip segments. In both *Young–Long* and *Old–Long*, the integrated *Proposed* specification performs better than both versions of the *Reference* setup. This remains only a relative improvement, since the resulting correlations are still modest in absolute terms. Moreover, the distance-distribution results do not indicate that long leisure travel is well reproduced overall in the simulation. The spatial-correlation results should therefore not be interpreted as evidence that the long-trip behaviour is fully captured. A more cautious interpretation is that, conditional on such trips being realised within the SIMBA environment, the *Proposed* specification appears somewhat better able than the *Reference* setup to direct them towards destinations that are attractive for long leisure trips. In this sense, the gain seems to lie more in the local spatial allocation of realised long trips than in a broader improvement of long-distance leisure mobility as such.

7 Discussion

This chapter brings together the main results of the thesis. Section 7.1 discusses the main findings, Section 7.2 reflects on the main limitations of the study, Section 7.3 outlines directions for future research, and Section 7.4 considers implications for practice and policy.

7.1 Discussion of main findings

This thesis addressed a common limitation of leisure destination-choice modelling, namely the use of simplified and weakly differentiated representations of destination attractivity. To address this gap, it developed a multidimensional zone-level leisure attractivity index designed to capture the broader territorial context of leisure destinations while remaining transparent, interpretable, and operationally tractable. The framework was then assessed from three complementary perspectives. First, whether such an attractivity formulation can be meaningfully defined and operationalised. Second, whether it helps reveal which destination-side dimensions matter for leisure destination choice across different travel contexts and population groups. Third, whether it adds explanatory, predictive, and operational value relative to simpler baseline formulations, including a first proof-of-concept integration into the SIMBA MOBi framework. The main findings are discussed below in relation to these three research questions.

RQ1 – Definition and operationalisation of leisure attractivity. The first research question asked how leisure attractivity $A_j(s)$ can be defined and operationalised at the traffic-zone level so as to capture the territorial context and heterogeneity of leisure opportunities. Overall, the findings suggest that this can be done in a richer way than through a small number of generic proxies. A richer operationalisation is possible when attractivity is treated as a multidimensional destination-side construct, in line with the literature (Crouch, 2011; Dwyer & Kim, 2003; Świdyńska & Witkowska-Dąbrowska, 2021). In this thesis, this was achieved by starting from recurrent leisure domains and destination-side dimensions identified in the literature (Section 3.2) and translating them into a finite set of observable zone-level indicators (Section 4.3). Concretely, this meant representing destinations through a broader set of leisure-relevant characteristics, including amenities and points of interest, density and diversity measures, outdoor and sport-related infrastructure, supporting infrastructure, and environmental features such as land-use context and access to blue spaces.

A central objective was to keep the formulation interpretable, transparent, and empirically traceable. For this reason, the selected indicators were chosen not only for their conceptual relevance, but also for their semantic clarity and limited redundancy. The index was built from indicators that could be linked to recognisable dimensions of leisure opportunities. For example, counts of restaurants, cafés, and bars can be linked to a gastronomic dimension, diversity measures to the variety of leisure offer, the length of hiking trails and skiing tracks to outdoor and active sport recreation, and lake shore length to outdoor recreation and the environmental appeal of destinations. Their semantic meaning is relatively clear and they show limited conceptual overlap, which supports the interpretability of the resulting formulation.

The proposed index should be understood not as a mechanically exhaustive inventory of leisure, but as a balanced and theory-guided representation of the main destination-side features. The formulation aimed to strike a balance between studies that emphasise specific domains of leisure and their distinct destination requirements (Gramsch-Calvo & Axhausen, 2025; Simma et al., 2001) and the need to avoid excessively large sets of variables that would become difficult to interpret and operationalise in practice (Biernacka et al., 2020; Krešić & Prebežac, 2011; Singh & Tiwari, 2016). In its final form, the thesis arrived at a formulation based on 13 variables (Section 5.1). While this remains manageable and interpretable, an even more compact specification could prove advantageous for operational application.

An important part of this contribution is the use of open spatial data, especially OpenStreetMap. OSM provides a rich and flexible source of geographically explicit information. Its use is also consistent with recent work showing the value of open-data-based indicators for enriching destination-side representations in travel demand modelling (Klinkhardt et al., 2021; Notelaers et al., 2024). At the same time, the thesis does not treat these data uncritically. Their plausibility was assessed explicitly, and the resulting indicators were retained with attention to completeness and internal consistency (Appendix D). In this sense, the use of OSM and other open spatial data contributes not only to a richer representation of leisure destinations, but also to a more transparent and largely reproducible attractivity formulation. Some dimensions, however, still require complementary sources.

The resulting index remains a partial representation of leisure attractivity rather than a direct measurement of it. More qualitative and subjective dimensions, such as destination loyalty, atmosphere, familiarity, or emotional attachment, are also important for destination choice (Leong et al., 2015; Moreira & Iao, 2014; Reitsamer et al., 2016; Yoon & Uysal, 2005). These remain, however, difficult to capture through spatial indicators alone. Overall, the findings suggest that leisure attractivity can be operationalised at the traffic-zone level as a multidimensional, interpretable, and reproducible destination-side representation. The proposed index captures the territorial context and heterogeneity of leisure opportunities, while remaining an indirect representation of leisure attractivity.

RQ2 – Segment-specific patterns of destination-side preferences. The second research question asked whether the proposed leisure attractivity index is able to reveal differences in destination-side preferences across travel contexts and population groups, and which destination-side variables are associated with leisure destination choice in each case. In this thesis, this question was addressed as a pragmatic first step through two ex-ante segmentation dimensions: age group and trip context. Concretely, this meant distinguishing between individuals aged under 65 and those aged 65 and over, and between shorter leisure travel embedded in everyday routines and longer day-trip-type leisure travel (Young–Short, Old–Short, Young–Long, Old–Long). Overall, the results suggest that the richer formulation provides an interpretable structure through which such heterogeneity can emerge empirically. Because the specification is built from transparent variables that can be linked to recognisable dimensions of leisure opportunities, it becomes possible to identify which destination-side features are positively or negatively associated with the estimated utility of choosing a destination zone in each behavioural context.

The clearest differences emerge across travel contexts rather than across age groups. The contrast between short and long leisure trips is strong and systematic. Across both age classes, several coefficients change sign or magnitude when moving from short to long trips. For example, the associations of sport-related points of interest, cultural points of interest, and the land-use mix index are positive in the short-trip models but negative in the long-trip models. By contrast, the positive associations of gastronomy-related points of interest, outdoor sport route length, and the leisure-offer variety index become stronger in the long-trip models. This indicates that similar destination environments are associated differently with destination utility depending on trip context. A plausible interpretation is that longer leisure trips are more often directed towards specialised excursion destinations rather than towards zones with a broader mix of urban leisure opportunities. In the Swiss context, this is consistent with alpine destinations, where hiking or skiing infrastructure and gastronomy jointly contribute to destination appeal. This pattern is behaviourally plausible and consistent with the literature showing, more generally, that leisure travel is not homogeneous across contexts: different leisure activities are associated with different travel distances (Spinney & Millward, 2013), and same-day and longer leisure travel can follow different behavioural logics (Scotti et al., 2024).

Age-related differences are present, but weaker than differences by travel context. Even so, the richer formulation still yields some meaningful insights. Sport-related POIs show a stronger positive association for younger individuals than for older ones, which is consistent with the exploratory data analysis, where sport was the trip-purpose category showing the clearest difference between the two age groups. Urban POIs density is positively associated with destination utility for younger individuals in the short-trip context, but negatively associated for older individuals. This may indicate that older individuals are less likely to choose destinations characterised by dense and highly urban settings, whereas younger individuals show a more positive association with such environments. Support amenities show a stronger positive association for older individuals in the short-trip context. These contrasts are not large, but they remain broadly plausible. Overall, this pattern is more consistent with the more cautious view, according to which age-related differences exist but remain limited (Buehler et al., 2024; Reece, 2004). More generally, it also supports the broader argument that heterogeneity in leisure-related behaviour cannot be reduced to age alone (Ohnmacht et al., 2009; Shoemaker, 1989).

At the same time, somewhat stronger differences between age groups might have been expected. A first explanation is that the age segmentation adopted here remains relatively coarse. In particular, the broad under-65 group likely combines individuals with quite different lifestyles, time constraints, and leisure preferences, which may blur age-related contrasts. The additional exercise reported in Appendix K explores a finer age segmentation, but the resulting differences still do not become consistently sharp across all cases, suggesting again that age alone may not be the main dimension along which leisure destination preferences vary. A second explanation is that some age-relevant dimensions may not be sufficiently captured in the present specification, such as place reputation, perceived quality, comfort, service capacity, or qualitative information on amenities. Finally, age-related differences may emerge more clearly at a finer spatial scale, not so much in the choice of the destination zone as in the choice of the specific destination within the zone, especially where the local leisure offer is internally diverse.

Across the four segment-specific models, some features remain repeatedly important, most notably gastronomy amenities, route-based outdoor infrastructure, and support amenities. Gastronomy appears as a broadly relevant component of leisure destination utility, in line with evidence that food-related opportunities contribute to destination appeal and travel experience (Sirkis et al., 2022). Route-based outdoor infrastructure also emerges as consistently relevant, consistent with work showing that trails and skiing opportunities can structure recreational attractivity (Molloy & Moeckel, 2017; Willibald et al., 2019). Support amenities, in turn, seem to capture a broader enabling dimension of destination utility, in line with evidence that service availability matters for how destinations are evaluated and used (Singh & Tiwari, 2016). Some dimensions, however, appear less influential than expected. The land use mix indicator plays a weaker role than expected, despite evidence that it can strengthen destination attractivity (Chen et al., 2024). Cultural indicators are also less prominent than expected, even though cultural amenities are linked to territorial attractivity and visitor appeal (Marrocu & Paci, 2013). In the present setting, these dimensions may therefore be more context-specific or only partially captured by the current specification.

Overall, the findings suggest that the proposed leisure attractivity index is useful for identifying which destination-side features are associated with leisure destination choice. It is also useful for showing how these associations vary across travel contexts and, more modestly, across age groups. Its strength lies not only in improving destination-side representation, but also in linking estimated effects to interpretable dimensions of leisure attractivity. In this sense, the index improves both the representation and the substantive interpretation of heterogeneity in leisure destination choice.

RQ3 – Added value relative to simpler baselines. The third research question asked to what extent a richer and segment-specific representation of leisure attractivity improves the explanatory and predictive performance of leisure destination choice models relative to simpler baselines (the SIMBA MOBi reference specification and an accessibility-only specification). The results suggest that the proposed specification adds explanatory and predictive value, although this added value is not equally strong across all considered segments and does not translate directly into equally clear improvements within the SIMBA MOBi system. In particular, the improvements are clearest for long leisure trips, where a richer destination-side representation appears most informative. The improvements are more modest for short trips, where accessibility already accounts for a large part of destination choice. Moreover, these gains emerge more clearly in the offline econometric assessment than in the proof-of-concept SIMBA MOBi integration. Here, the interpretation remains constrained by the absence of full system recalibration and by the approximate operational translation of the empirical segmentation.

From the econometric perspective, the proposed specification performs better than the simpler benchmark models in the out-of-sample evaluation in terms of individual-level probability fit, ranking metrics, and aggregate destination-share reproduction (see Table 10). By contrast, average predicted travel distance is reproduced slightly more closely by the simpler benchmark specifications. For short leisure trips, the gains relative to the simpler baseline specifications are limited. Accessibility alone already explains a substantial part of destination choice in these segments, suggesting that proximity and general accessibility constraints remain the dominant drivers of daily routine behaviour. Consistently, the out-of-sample differences in McFadden pseudo-R^2, negative log-likelihood, and the correlations between observed and model-implied destination shares remain relatively small in the short-trip segments. The picture is different for long leisure trips. Although these segments are more difficult to predict overall, as indicated by their weaker performance across the main evaluation metrics compared with short trips, they are the ones for which the richer specification delivers the clearest improvement over both baseline models. In this sense, the proposed attractivity structure captures the spatial logic of leisure destination choice more successfully in leisure trips that are less embedded in everyday routines. This is substantively important, because these trips are precisely the context in which accessibility alone becomes less sufficient and destination-side characteristics matter more. The richer formulation therefore appears most useful where leisure destination choice is less governed by routine proximity constraints.

A further result concerns age segmentation specifically. Overall, the gains from estimating the model separately by age group remain smaller than those observed for trip context. The short-trip results still reveal some age-related differences. Compared with a pooled estimation without age segmentation, model performance improves for the Old–Short segment, while it remains similar or slightly weaker for the Young–Short segment. This suggests that the pooled specification partly masks a more distinct behavioural profile for the older group, while instead approximating behaviour that is closer to the overall average. This may also be related to sample composition, since the smaller number of observations for the 65+ group makes their specific behavioural pattern less likely to emerge clearly. Even so, the gain for the older short-trip segment remains limited in magnitude, which is consistent with the broader finding that age-related heterogeneity is present but weaker than heterogeneity by trip context. The finer age-segmentation exercise reported in Appendix K points in the same direction. Some additional variation emerges, with comparatively better performance for the youngest and older age classes, while the middle-aged groups remain closer to the pooled pattern. This suggests that, especially for the young segment in the main analysis, other dimensions not explicitly modelled here may further help explain leisure destination choice.

In the SIMBA MOBi integration, the gains observed in the offline econometric evaluation do not carry over to the same extent into the operational simulation results. The exercise should therefore be understood as a proof of concept. Its main contribution is to show that a richer and segment-specific leisure destination utility can be embedded in a country-wide activity-based simulation model without destabilising the broader system. This demonstrates that the proposed specification is not only estimable offline, but also operationally implementable within a large-scale modelling framework. This, in turn, creates a basis for future scenario testing within SIMBA MOBi or other modelling environments. The segment-specific specification makes it possible to examine how demographic ageing may affect leisure destination choice, since changes in population composition can be linked to different destination-choice responses across age groups. At the same time, because the proposed attractivity term is expressed through observable territorial characteristics, the framework can also be used to analyse how different population segments may respond to changes in the spatial configuration of leisure opportunities, services, or environmental amenities. This is particularly relevant given the expected growing importance of leisure mobility in future travel demand. Moreover, if individuals have more discretionary time available, longer leisure trips that occupy a substantial part of the day may also become more important.

At the same time, the integration results must be interpreted with caution for three main reasons. First, the implemented specification corresponded to a slightly earlier version of the estimated leisure utility rather than to the final version reported in this thesis. Second, the surrounding SIMBA MOBi system was intentionally left unchanged, so the new destination utility was introduced without a broader recalibration of the model environment in which it operates. Third, the operational assignment of the long-trip utility relied on a proxy rule rather than on a direct equivalent of the empirical segmentation used in the offline analysis, which contributed to a strongly unbalanced segment distribution in the simulated outputs.

Against this background, the comparison between the two simulation specifications (the SIMBA MOBi reference specification and the proposed specification) focuses primarily on two aspects. First, the realised travel-distance distributions generated by the simulation. Second, the aggregate spatial destination patterns, measured through destination shares and compared with the observed leisure trips in the Swiss Mobility and Transport Microcensus (MTMC). For short trips, the proposed specification yields only limited changes in the simulated distance distributions. In some cases, the results move slightly towards the MTMC reference, but the improvement remains small. The observed spatial destination patterns are reproduced more closely by the reference specification. For long trips, distance reproduction remains clearly poor under both specifications. A plausible explanation is that the long-trip utility is embedded in a simulation environment still governed by strong daily scheduling, time-budget, and feasibility constraints. These constraints may limit the emergence of genuinely long leisure trips even when the destination utility itself points towards more distant alternatives. At the same time, the long-trip results are not entirely discouraging. Although the distance patterns remain problematic, the spatial destination patterns for long trips are reproduced somewhat better under the proposed specification than under the reference setup. This suggests that, within the simulation environment, the richer destination-side structure still retains some value in directing realised long trips towards more plausible destinations. This is an encouraging indication, especially given that longer leisure trips had been identified by SIMBA MOBi experts as a relevant blind spot in the current operational representation.

Overall, the results suggest that a richer and segment-specific representation of leisure attractivity improves the explanatory and predictive performance of leisure destination choice models relative to simpler baseline specifications. This improvement is clearest in the offline econometric evaluation, especially for leisure trips that are less embedded in everyday routines, where destination-side heterogeneity appears to matter more strongly. At the same time, the findings also show that these gains do not transfer automatically into the operational SIMBA MOBi environment. Even so, the proof-of-concept integration still demonstrates the practical feasibility of embedding such a richer and segment-specific leisure destination utility within a country-wide activity-based simulation framework.

Overall synthesis. Overall, this thesis addresses a practical gap in leisure destination-choice modelling. On the one hand, operational travel demand models often rely on simplified destination-attractivity terms, typically based on a small number of aggregate zonal proxies such as resident population or employment. On the other hand, richer representations of destination choice can quickly become difficult to interpret, estimate, and integrate into large-scale operational simulation systems. Against this background, the thesis develops a richer but still interpretable multidimensional index of leisure destination attractivity. The aim is not only to represent the diversity of destination-side characteristics that are relevant for leisure travel, but also to assess whether such a representation helps reveal interpretable differences across travel contexts and demographic groups.

The main contribution of the thesis lies in showing that a more explicit, leisure-specific, and behaviourally differentiated destination-side structure can be constructed, estimated, validated, and shown to improve predictive performance, while remaining transparent and operationally tractable. Because the formulation is built from explicit and semantically meaningful variables, the estimated effects can be linked to recognisable dimensions of leisure attractivity and interpreted more substantively than in simpler aggregate specifications. The framework was also benchmarked against simpler specifications and subjected to a broad set of robustness checks. The main findings and performance patterns proved generally stable across alternative modelling choices (see Section 6.3). In this sense, the thesis offers a middle ground between overly coarse attractivity terms and specifications that become too complex for applied use. More broadly, it provides a structured basis for better understanding and modelling leisure mobility through destination choice, and it shows that such a representation can also be translated into a first operational proof of concept within a larger simulation system.

7.2 Limitations

For clarity, the limitations are grouped into five families: data limitations, methodological limitations, design and specification choices, computational and operational limitations, and broader conceptual limitations.

Data limitations. The Microcensus is one of the most suitable national sources for observing mobility behaviour, but it remains a survey-based sample and is therefore affected by the usual limitations of survey data, including sampling noise, possible underrepresentation of some groups, and limited observations in some segments. In this thesis, this is reflected most clearly in the low number of observations for the Old–Long segment. The reason is that the survey module used to record longer trips was administered only to one third of participants. For an age group that already represents a smaller share of the sample, this further exacerbates the scarcity of observations.

Most destination-side indicators are derived from OpenStreetMap and related spatial data. This supports openness and reproducibility, but also introduces limitations related to completeness, imperfect tagging, spatial unevenness, and missing attribute information. In this thesis, the resulting indicators mainly capture relatively static aspects of the quantity and spatial distribution of leisure opportunities. By contrast, potentially relevant dimensions such as quality, capacity, and temporal availability could not be represented satisfactorily in the present study.

The accessibility term EMU_{ij} is coherent with the SIMBA MOBi framework, but it also reflects the specific modelling logic and assumptions of that system. In this sense, it is not a fully generic accessibility measure, but one that is partly shaped by the operational context in which it was originally defined. At the same time, the robustness analysis suggests that replacing EMU_{ij} with simpler accessibility measures that are not tied to the SIMBA MOBi framework does not materially alter the main results.

In addition, the three main data components are not perfectly aligned in time: the observed trips come from 2015 and 2021, the accessibility from 2023, and the supply-side indicators reflect conditions from roughly 2024–2026. This temporal mismatch is difficult to avoid fully and should be kept in mind.

Methodological limitations. The proposed attractivity index is specified as a linear combination of thirteen destination-side variables. This improves interpretability, but it cannot fully capture non-linearities, threshold effects, or more complex interactions between destination features.

Within each segment, behaviour is still treated as homogeneous. The segmentation strategy adopted in this thesis allows part of the heterogeneity in leisure destination choice to be examined explicitly, but it cannot capture all remaining variation within each group. In particular, the model does not represent richer forms of unobserved heterogeneity, such as random taste variation or repeated choices by the same individual. In addition, the adopted multinomial logit structure inherits the usual IID assumption across alternatives, which may be restrictive when destinations are spatially close, functionally similar, or jointly perceived (Simma et al., 2001). More generally, unobserved influences may also vary across spatial contexts, which the present specification does not model explicitly (Simma et al., 2001).

The estimated coefficients should not be interpreted causally. They reflect conditional associations within the adopted specification and may still be affected by omitted variables or residual collinearity. More generally, model evaluation in very large choice sets remains difficult: only realised choices are observed, not the individual's true considered set. The resulting estimates should therefore be interpreted as informative about how destination-side characteristics are associated with leisure destination choice.

Design and specification choices. The segmentation strategy by age group and trip type is behaviourally motivated and well suited to the aims of this thesis. However, it may not be the segmentation that reveals behavioural differences most sharply. Other dimensions, such as income, household composition, work schedule, seasonality, or weather, could also matter. At the same time, introducing additional segmentation dimensions would have substantially increased the complexity of the analysis.

The age split is especially simplified. In particular, the 6–64 group remains highly heterogeneous, and the sensitivity analysis with finer age groups suggests that age-related heterogeneity is only partly captured by the binary split. At the same time, the finer age segmentation does not indicate a uniformly large additional gain in explanatory or predictive value across all groups. The distinction between short and long leisure trips is operationally useful and closely aligned with the structure of the Microcensus. This makes the segmentation substantively meaningful, but also partly dependent on the underlying survey design. Alternative definitions based on time or distance thresholds could therefore have led to somewhat different results.

A further design choice concerns the spatial scope of the analysis. Estimating the model at the level of the whole country makes it possible to retain a unified and operationally relevant framework, but it also means that the estimated relationships are inferred across very different territorial contexts. Urban, peri-urban, and rural areas, for example, may differ substantially in both the structure of leisure opportunities and the way destination choice is formed.

Destination choice is modelled at the zonal level rather than at the level of specific facilities or micro-locations. This is consistent with operational travel demand modelling practice, but it inevitably abstracts from within-zone heterogeneity. More generally, the destination-side variables describe the broader context of a zone rather than the exact reason why an individual chose a specific destination within it. The estimated effects should therefore be interpreted as associations between destination environments and observed choices at the zonal level.

A related design simplification is that leisure travel is treated in an aggregated way. The model remains largely agnostic about the specific underlying purpose of a given leisure trip, such as gastronomy, walking, cultural activities, or social visits. This abstraction is consistent with the objective of modelling leisure destination choice as a whole, but it also means that some activity-specific differences in destination preferences are not explicitly represented.

Computational and operational limitations. The destination-choice problem considered in this thesis is computationally demanding because it involves a very large choice set. For this reason, estimation relies on sampled choice sets rather than on the full set of available destinations. Although the robustness checks suggest that this does not materially alter the main conclusions, it remains a computational limitation of the empirical strategy.

A further set of limitations concerns the operational integration into SIMBA MOBi. First, for computational reasons, the version of the leisure utility that was integrated into SIMBA MOBi corresponded to an earlier specification than the final one reported in this thesis. Although the two versions remain broadly similar in their implied attractivity structure, this discrepancy calls for additional caution in interpreting the operational results. Second, the simulation experiment could only be implemented on a reduced agent sample (10%). Third, the empirical short–long segmentation could only be translated approximately into the simulation environment, through a proxy rule based on agents' daily plans. This first operational approximation contributed to an imbalanced segment distribution in the simulated outputs, including too few trips in the Old–Short segment. Fourth, the new leisure utility was inserted into the existing SIMBA environment without a broader recalibration of the surrounding model system. This helps isolate the contribution of the new specification, but it also means that the integrated utility could not fully express the potential indicated by the offline econometric analysis.

A final operational limitation concerns the temporal scope of the simulation itself. SIMBA MOBi represents an average weekday, which is not necessarily the most suitable context for all forms of leisure mobility. This is particularly relevant for longer excursion-type leisure trips, which may be more frequent or behaviourally more important on weekends or other non-routine days.

Broader conceptual limitations. Many relevant dimensions of destination choice, such as destination image, place attachment, loyalty, symbolic meaning, reputation, or the role of specific events, are difficult to observe and even more difficult to quantify consistently. More generally, destination attractivity is not a directly observable quantity for which a universal measure exists. It must instead be approximated through a set of theoretical, empirical, and operational choices concerning which dimensions to include, how to represent them, and how far the resulting formulation can be considered sufficiently representative. For this reason, any attractivity index, including the one proposed here, necessarily remains a partial and constructed representation rather than a complete measurement.

A further fundamental limitation concerns external validity. The general workflow developed in this thesis is transferable, but the estimated index and coefficients are specific to the Swiss case study and to the data sources used here. They should therefore not be assumed to apply directly in other geographical or institutional contexts without re-estimation. In this sense, what is transferable is primarily the modelling framework and its logic, rather than the exact estimated attractivity structure itself.

7.3 Future work

Although not all of the limitations discussed above can be fully resolved, they point to several promising directions for further work. These concern the empirical basis of the analysis, the modelling framework, the segmentation strategy, the operational integration into SIMBA MOBi and scenario analysis, and the broader conceptualisation of leisure attractivity.

A first natural extension concerns the empirical basis of the study. The next wave of the Swiss Mobility and Transport Microcensus, expected in 2027, would offer an important opportunity to re-estimate the framework with improved temporal consistency between observed trips, supply-side indicators, and accessibility inputs. Another direction would be to enrich the supply-side representation with information on dimensions that could not be captured satisfactorily in the present study, such as quality, popularity, capacity, or effective use. This could be explored by integrating additional data sources where available, such as online ratings or specialised surveys on place quality. This would move the attractivity index beyond a primarily quantity-oriented representation of leisure opportunities. At the same time, integrating such measures may also involve a trade-off, since it can reduce interpretability and the direct link between the attractivity formulation and observable territorial structure.

A second direction concerns the modelling framework itself. The specification adopted in this thesis is intentionally relatively simple, transparent, and interpretable. Future work could therefore test richer discrete-choice formulations, for example mixed logit or error-components specifications (Train, 2009). At the same time, such extensions should be considered with caution, since greater flexibility may come at the cost of reduced interpretability and operational tractability.

A third avenue concerns segmentation and behavioural scope more broadly. Because the framework is transferable and sufficiently flexible, it can be extended to explore a wider range of behavioural contexts than those considered in this thesis. Future work could apply the same modelling logic to other sociodemographic segmentations, to different contextual partitions, or to more local spatial scales where destination environments may be more internally coherent. Another promising direction would be to estimate separate attractivity formulations for specific types of leisure travel. In this way, the framework could help identify in which contexts destination-side preferences are most clearly differentiated and most meaningfully expressed.

A fourth and particularly important direction concerns the operational integration into SIMBA MOBi. A first priority for future work is to implement the final corrected leisure utility reported in this thesis. Beyond that, the proposed specification should be better calibrated within the SIMBA workflow, so that its contribution can be assessed jointly with the surrounding behavioural modules rather than in partial isolation. Future work should also make use of one of the main strengths of the proposed framework, namely its suitability for scenario testing. Because the utility is both segment-specific and based on observable territorial characteristics, it offers a basis for simulating how leisure destination choice may respond to future demographic ageing, shifts in the balance between population groups, or changes in the spatial distribution of leisure opportunities and amenities. In this sense, the framework could support analyses that go beyond what is possible with the current more aggregated specification.

A separate avenue for future research is to test the framework beyond the Swiss case study. Applying the same workflow in other national or regional contexts would provide the clearest assessment of its broader transferability. Such extensions would, however, depend on the availability of sufficiently rich travel-survey, accessibility, and spatial supply data in the respective settings.

7.4 Implications for practice and policy

The contribution of this thesis is not only methodological, but also practical and policy-relevant. Leisure mobility already accounts for a substantial share of total travel demand and is expected to become even more important in the future. This broader relevance is reinforced by demographic ageing and by ongoing changes in travel behaviour, which may make mobility demand less regular and less predictable than in the past. As emphasised in SBB's Strategy 2030, one of the key challenges for the future will be to respond to increasingly diverse and flexible travel patterns through a more customer-oriented and adaptable transport system (SBB, 2021). In this context, improving the behavioural representation of leisure mobility is not merely a modelling refinement, but also a relevant step for transport planning and policy analysis.

A central limitation of many current operational travel demand models is that leisure travel, and in particular destination attractivity, is often represented only in a simplified way, while population heterogeneity is not represented in a behaviourally meaningful manner. This makes it more difficult to analyse how destination preferences differ across travel contexts and population groups, and to assess how these differences may evolve under future demographic scenarios. Such limitations are particularly relevant in a context where transport operators are increasingly confronted with leisure demand that contributes to irregular peaks and more complex temporal patterns outside traditional commuting periods. The framework proposed in this thesis is relevant in this respect because it provides a richer and more interpretable representation of leisure destination attractivity, while also making it possible to study behavioural differences across population groups and travel contexts.

This is particularly valuable for scenario analysis. Because the proposed attractivity term is expressed through observable territorial characteristics, it can support the assessment of how changes in the built environment may affect future leisure demand in specific areas. This includes, for example, the opening of new restaurants or cultural venues, the redevelopment of public spaces, or broader changes in the spatial configuration of local leisure opportunities. Since gastronomy-related amenities show a positive association with leisure destination utility across the estimated models, one concrete application would be to test, in scenario form, how changes in such amenities in a given area could affect its predicted attractivity and destination shares. In this sense, mobility policy and territorial policy are closely linked, since accessibility and the spatial distribution of leisure opportunities jointly shape destination choice. The framework is equally relevant for transport scenarios, for example when evaluating how changes in infrastructure or in public-transport service levels may affect leisure mobility patterns. This matters in particular because leisure trips are often less routine-driven than commuting trips and may therefore be more sensitive to flexibility, service quality, and overall travel experience. A richer destination-choice representation can therefore help identify areas that may become relevant for leisure demand even if they are less central from a commuting perspective, thereby informing service planning beyond traditional peak-period travel. The segment-specific specification also has value for long-term demand projections in a context of demographic ageing. In the case of this thesis, applying age-differentiated destination-choice parameters to future synthetic populations could help represent more realistically how leisure demand changes as the relative size of different age groups evolves over time. For example, the weaker association between older individuals and sport-related amenities suggests that a growing older population could shift the relative importance of different leisure destination types, although such implications would need to be tested through explicit scenarios rather than inferred directly from the coefficients alone. A similar point applies to longer leisure trips, which emerged in this thesis as more behaviourally distinctive and more difficult to represent within the current operational setting, but which may also gain importance in the future. Scenario applications could therefore examine how leisure destination patterns change when more individuals have longer time windows available for discretionary travel, and which areas would receive higher predicted demand under such conditions.

More generally, the same framework could be used to explore other forms of behavioural change, provided that they can be linked either to differences in segment composition or to changes in the territorial structure of leisure opportunities. Overall, these examples show that a better representation of leisure mobility, achieved through a more behaviourally meaningful destination-attractivity term, is relevant not only for model estimation but also for policy evaluation, transport planning, and forward-looking scenario analysis.

8 Conclusion

This thesis set out to improve the representation of leisure mobility by refining one important but often under-specified component of travel demand models, namely the destination-side attractivity term, often referred to as the size term. In a context in which leisure mobility is expected to become increasingly important, partly in connection with demographic ageing, improving its representation is both relevant and timely. At the same time, leisure mobility remains a complex phenomenon, shaped by many interacting features and not fully reducible to a single modelling component. The objective of this thesis was to develop a richer zonal attractivity term, one that is multidimensional and capable of reflecting a broader set of destination-side characteristics, while still remaining sufficiently interpretable and compact to be useful in applied and potentially operational contexts.

To this end, the thesis developed a structured destination-side attractivity formulation and assessed whether it could improve model performance while also revealing differences in destination-choice behaviour across travel contexts and demographic groups. As a pragmatic first step, the empirical analysis focused on simple segmentation by age group (6–64 vs. 65+) and travel context (short everyday leisure vs. long day-trip leisure), yielding four segments: Young–Short, Old–Short, Young–Long, and Old–Long. This segmentation was not intended to be exhaustive, but rather to provide a meaningful testing ground for evaluating whether a richer attractivity formulation can make heterogeneity more visible and interpretable. The proposed index was estimated and validated separately for these four segments within a multinomial logit destination-choice framework, using observed leisure destination choices extracted from the national travel diary survey, the *Swiss Mobility and Transport Microcensus*. Model performance was assessed through out-of-sample evaluation using individual-level probability-fit measures, ranking metrics, and aggregate destination-share comparisons, while the estimated coefficients were interpreted in substantive terms to understand which destination-side features were most closely associated with leisure destination choice.

The results show that a richer, yet still compact and interpretable, representation of destination attractivity is possible and useful. The proposed specification adds value in both explanatory and predictive terms, while also improving the substantive interpretation of leisure destination choice. More specifically, it helps reveal which territorial characteristics appear to matter most, which dimensions remain important across segments, and where meaningful differences emerge between the considered behavioural contexts. In the present case, the richer formulation was more effective in explaining differences between short and long leisure trips than differences between age groups. This is an important result, since it suggests that richer destination-side modelling can help identify not only whether heterogeneity matters, but also along which dimensions it matters most clearly.

The thesis also shows that transferring these gains into an operational model environment is not automatic. The proof-of-concept integration into SIMBA MOBi confirmed that such an attractivity formulation can be embedded into a real, country-wide operational simulation framework without destabilising the broader system. This shows that the proposed specification is not only theoretically meaningful and econometrically estimable, but also practically implementable in an applied modelling environment. However, the integration also showed that operational benefits require more than replacing a single utility component. The proposed specification needs to be adapted and recalibrated within the broader activity-based simulation environment before its full benefits can be realised.

Overall, this thesis does not resolve the full complexity of leisure mobility, but it shows that there is meaningful space between highly simplified attractivity terms and models that become too complex to interpret or implement. Within that space, it is possible to construct a destination-side representation that is richer, empirically grounded, reproducible, and still manageable in practice. In this sense, the thesis contributes a useful framework both for the scientific analysis of leisure destination choice and for the future development of more behaviourally informed modelling tools in applied planning contexts. It also provides a basis for more policy-relevant scenario analysis, particularly for assessing how demographic ageing and changes in the spatial distribution of leisure opportunities may shape future leisure mobility patterns.

References

Amt für Mobilität, Kanton Zürich. (2025). *Zürcher Gesamtverkehrsmodell (GVM-ZH)*. Amt für Mobilität, Kanton Zürich. Available at: https://www.zh.ch/de/mobilitaet/gesamtverkehrsplanung/verkehrsgrundlagen/gesamtverkehrsmodell.html [Accessed 10.11.2025].

Amt für Statistik Liechtenstein. (2025). Liechtenstein in Zahlen 2025. Amt für Statistik Liechtenstein. Available at: https://www.statistikportal.li/statistikportal/publications/102-liechtenstein-in-zahlen/2025/01/1/102.2025.01.1_02_liechtenstein-in-zahlen-2025.pdf [Accessed 28.01.2026].

Anda, C., Erath, A., & Fourie, P. J. (2017). Transport modelling in the age of big data. *International Journal of Urban Sciences*, *21*(sup1), 19–42.

Anselin, L. (1995). Local indicators of spatial association–LISA. *Geographical Analysis*, *27*(2), 93–115.

Anselin, L. (2022). Spatial econometrics. In S. J. Rey & R. S. Franklin (Eds.), *Handbook of spatial analysis in the social sciences* (pp. 101–122). Edward Elgar Publishing.

ARE. (2017). NPVM 2016: Zonenstruktur und Verkehrsnetze. Schlussbericht. Bundesamt für Raumentwicklung (ARE). Available at: https://www.are.admin.ch/dam/de/sd-web/NSr6UDjeuecu/beicht-npvm-zonen-netze-2016%20.pdf [Accessed 28.01.2026].

ARE. (2020). Modelletablierung Nationales Personenverkehrsmodell (NPVM) 2017. Schlussbericht. Bundesamt für Raumentwicklung (ARE). Available at: https://www.are.admin.ch/dam/de/sd-web/EBN5gUfmwYva/NPVM_2017_Schlussbericht.pdf [Accessed 14.11.2025].

ARE. (2020). Nationales Personenverkehrsmodell NPVM 2017. Grundlagen. Bundesamt für Raumentwicklung (ARE). Available at: https://www.are.admin.ch/dam/de/sd-web/4wxv-7k5YbPh/npvm2017.pdf [Accessed 14.11.2025].

ARE. (2021). Schweizerische Verkehrsperspektiven 2050. Schlussbericht. Bundesamt für Raumentwicklung (ARE). Available at: https://www.digitaleschweiz.ch/wp-content/uploads/2021/11/verkehrsperspektiven-schlussbericht.pdf [Accessed 23.10.2025].

ARE. (2022). *Transport Outlook 2050: Developments in passenger transport*. Bundesamt für Raumentwicklung (ARE). Available at: https://www.are.admin.ch/en/developments-in-passenger-transport [Accessed 23.10.2025].

ARE. (2025). *Nationales Personenverkehrsmodell (NPVM)*. Bundesamt für Raumentwicklung (ARE). Available at: https://www.are.admin.ch/de/npvm [Accessed 14.11.2025].

ARE. (2026). Nationales Personenverkehrsmodell (NPVM): Aktualisierung 2023. Schlussbericht. Bundesamt für Raumentwicklung (ARE). Available at: https://www.are.admin.ch/dam/de/sd-web/4KquT1YIdGEZ/npvm2023.pdf [Accessed 03.03.2026].

ARE & BFS. (2021). *Mikrozensus Mobilität und Verkehr (MZMV)*. Bundesamt für Raumentwicklung (ARE) und Bundesamt für Statistik (BFS). Available at: https://www.are.admin.ch/de/mzmv [Accessed 23.10.2025].

ARE & BFS. (2023). *Mikrozensus Mobilität und Verkehr 2021: Kurzversion Fragebogen*. Bundesamt für Raumentwicklung (ARE) und Bundesamt für Statistik (BFS). Available at: https://www.bfs.admin.ch/bfs/rm/home/statisticas/catalogs-bancas-datas/publicaziuns.assetdetail.24845295.html [Accessed 23.10.2025].

ARE & SBB. (2024). Synthetische Population 2022. Modellierung mit dem Flächennutzungsmodell FaLC. Schlussbericht. Bundesamt für Raumentwicklung (ARE) und Schweizerische Bundesbahnen (SBB). Available at: https://www.are.admin.ch/dam/de/sd-web/WTmt7B7TDBf2/Schlussbericht_Synpop2022_Publikation.pdf [Accessed 14.11.2025].

ARE & UNIL. (2018). La mobilité de loisirs en Suisse: État des lieux et évolution 2005–2015. Analyse du microrecensement mobilité et transport. Office fédéral du développement territorial (ARE) et Université de Lausanne (UNIL). Available at: https://www.are.admin.ch/dam/de/sd-web/CKusiVHb5Llt/rapport_mobilite_de_loisirs.pdf [Accessed 28.01.2026].

Arentze, T. A. (2015). Individuals' social preferences in joint activity location choice: A negotiation model and empirical evidence. *Journal of Transport Geography*, *48*, 76–84.

Armellini, M. G., Banse Bueno, O. A., Bieker-Walz, L., Erdmann, J., Flötteröd, Y. P., & Rummel, J. (2021). Brunswick simulation scenario for virtual-stops based DRT services with SUMO. *In Proceedings of the 10th International Congress on Transportation Research*.

Axhausen, K. W., & Gärling, T. (1992). Activity-based approaches to travel analysis: Conceptual frameworks, models, and research problems. *Transport Reviews*, *12*(4), 323–341.

BAFU. (2025). *Swiss National Park and parks of national importance (Zones)*. Bundesamt für Umwelt (BAFU). Available at: https://data.geo.admin.ch/browser/index.html#/collections/ch.bafu.schutzgebiete-paerke_nationaler_bedeutung/items/schutzgebiete-paerke_nationaler_bedeutung [Accessed 28.01.2026].

Balmer, M., Nagel, K., & Raney, B. (2004). Large-scale multi-agent simulations for transportation applications. *Journal of Intelligent Transportation Systems*, *8*(4), 205–221.

Barceló, J. (2010). *Fundamentals of Traffic Simulation* (Vol. 145, p. 439). New York: Springer.

Bastarianto, F. F., Hancock, T. O., Choudhury, C. F., & Manley, E. (2023). Agent-based models in urban transportation: Review, challenges, and opportunities. *European Transport Research Review*, *15*(1), 19.

Becker, O., Börger, T., & Meyerhoff, J. (2025). Sampling of alternatives in spatial decision contexts with logit and logit mixture models: Simulation and application to freshwater recreation in Germany. *Resource and Energy Economics*, 101535.

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand* (Vol. 9). Cambridge, MA: MIT Press.

Ben-Akiva, M. E. (1973). *Structure of Passenger Travel Demand Models*. PhD thesis. Massachussets Institute of Technology (MIT).

BFS. (2008). *NOGA 2008: Allgemeine Systematik der Wirtschaftszweige. Anwendung zur Unterstützung der Codierung*. Bundesamt für Statistik (BFS). Available at: https://www.kubb-tool.bfs.admin.ch/de/noga/2008 [Accessed 03.02.2026].

BFS. (2010). Szenarien zur Bevölkerungsentwicklung der Schweiz 2010–2060. Bundesamt für Statistik (BFS). Available at: https://avenir50plus.ch/avplus50/wp-content/uploads/2016/12/Bevolkerungsentwicklung.pdf [Accessed 28.01.2026].

BFS. (2019). Bodennutzung. Bodennutzungsnomenklatur NOLU04. Bundesamt für Statistik (BFS). Available at: https://www.bfs.admin.ch/bfs/de/home/statistiken/raum-umwelt/nomenklaturen/arealstatistik/nolu2004.assetdetail.11087252.html [Accessed 28.01.2026].

BFS. (2023). *Statistik der Unternehmensstruktur (STATENT), Beschäftigte und Arbeitsstätten: Geodaten 2023*. Bundesamt für Statistik (BFS). Available at: https://www.bfs.admin.ch/bfs/de/home/dienstleistungen/geostat/geodaten-bundesstatistik/arbeitsstaetten-beschaeftigung/statistik-unternehmensstruktur-statent-ab-2011.assetdetail.36073031.html [Accessed 15.02.2026].

BFS. (2024). *Statistik der Bevölkerung und Haushalte (STATPOP), Geodaten 2024*. Bundesamt für Statistik (BFS). Available at: https://www.bfs.admin.ch/bfs/de/home/dienstleistungen/geostat/geodaten-bundesstatistik/gebaeude-wohnungen-haushalte-personen/bevoelkerung-haushalte-ab-2010.assetdetail.36171301.html [Accessed 15.02.2026].

BFS. (2025). *Ageing of the population*. Available at: https://www.bfs.admin.ch/bfs/en/home/statistics/population/ageing.html [Accessed 23.10.2025].

BFS. (2025). *Land use statistics based on the nomenclature NOAS04*. Bundesamt für Statistik (BFS). Available at: https://data.geo.admin.ch/browser/index.html#/collections/ch.bfs.arealstatistik/items/arealstatistik?.language=de-CH [Accessed 15.02.2026].

Bhat, C. R., & Koppelman, F. S. (1999). Activity-based modeling of travel demand. In R. W. Hall (Ed.), *Handbook of transportation science* (pp. 35–61). Springer US.

Biernacka, M., Kronenberg, J., & Łaszkiewicz, E. (2020). An integrated system of monitoring the availability, accessibility and attractiveness of urban parks and green squares. *Applied Geography*, *116*, 102152.

Buehler, R., Pucher, J., Wittwer, R., & Gerike, R. (2024). Trends and determinants of the mobility of older adults in the USA and Germany, 2001–2017. *Transportation Research Part A: Policy and Practice*, *183*, 104065.

Caicedo-Barreth, A., Santos Pavón, E., & Lima Santos, L. (2020). Competitiveness of Guayaquil towards bleisure tourism. *European Journal of Tourism, Hospitality and Recreation*, *10*(2), 118–133.

Chen, Y., Wang, B., Huang, J., Gao, H., & Shu, X. (2024). Urban physical environments promoting active leisure travel: An empirical study using crowdsourced GPS tracks and geographic big data from multiple sources. *Land*, *13*(5), 589.

Christensen, K., Doblhammer, G., Rau, R., & Vaupel, J. W. (2009). Ageing populations: The challenges ahead. *The Lancet*, *374*(9696), 1196–1208.

Chu, Z., Cheng, L., & Chen, H. (2012). A review of activity-based travel demand modeling. *CICTP 2012: Multimodal Transportation Systems — Convenient, Safe, Cost-Effective, Efficient*, 48–59.

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

Correia, A., Kozak, M., & Ferradeira, J. (2013). From tourist motivations to tourist satisfaction. *International Journal of Culture, Tourism and Hospitality Research*, *7*(4), 411–424.

Crompton, J. L. (1979). Motivations for pleasure vacation. *Annals of Tourism Research*, *6*(4), 408–424.

Crouch, G. I. (2011). Destination competitiveness: An analysis of determinant attributes. *Journal of Travel Research*, *50*(1), 27–45.

de Luca, S., & Cantarella, G. E. (2016). Validation and comparison of choice models. In *Travel demand management and road user pricing* (pp. 37–58). Routledge.

Dekker, T., Bansal, P., & Huo, J. (2025). Revisiting McFadden's correction factor for sampling of alternatives in multinomial logit and mixed multinomial logit models. *Transportation Research Part B: Methodological*, *192*, 103129.

Deng, M., & Athanasopoulos, G. (2011). Modelling Australian domestic and international inbound travel: A spatial–temporal approach. *Tourism Management*, *32*(5), 1075–1084.

Domencich, T. A., & McFadden, D. (1975). *Urban Travel Demand: A Behavioral Analysis*. Amsterdam: North-Holland.

Dong, X., Ben-Akiva, M. E., Bowman, J. L., & Walker, J. L. (2006). Moving from trip-based to activity-based measures of accessibility. *Transportation Research Part A: Policy and Practice*, *40*(2), 163–180.

Dwyer, L., & Kim, C. (2003). Destination competitiveness: Determinants and indicators. *Current Issues in Tourism*, *6*(5), 369–414.

Friedman, J. H., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*, 1–22.

Ghasemi, A., Miandoabchi, E., & Soroushnia, S. (2021). The attractiveness of seaport-based transport corridors: An integrated approach based on scenario planning and gravity models. *Maritime Economics & Logistics*, *23*(3), 522–547.

Ghodmare, S. D., & Yadav, G. (2021). Transportation planning using conventional four stage modeling: An attempt for identification of problems in a transportation system. *International Journal of Scientific Research in Science and Technology*, *8*(4), 556–567.

GLAMOS. (2016). *SCHWEIZER GLETSCHER*. Glacier Monitoring Switzerland (GLAMOS). Available at: https://www.glamos.ch/downloads#inventories/A10g-05 [Accessed 15.02.2026].

Goulias, K. G. (2018). Travel behavior models. In D. R. Montello (Ed.), *Handbook of behavioral and cognitive geography* (pp. 56–73). Edward Elgar Publishing.

Gramsch-Calvo, B., & Axhausen, K. W. (2025). The importance of the social environment on leisure destination choice: A mixed multinomial analysis of homophilic preferences. *Environment and Planning B: Urban Analytics and City Science*, *52*(7), 1650–1668.

Guevara, C. A., & Ben-Akiva, M. E. (2013). Sampling of alternatives in logit mixture models. *Transportation Research Part B: Methodological*, *58*, 185–198.

Guzmán, L. A., Cantillo-García, V. A., Oviedo, D., & Arellana, J. (2023). How much is accessibility worth? Utility-based accessibility to evaluate transport policies. *Journal of Transport Geography*, *112*, 103683.

Gwind S.r.l. (2023). *Popolazione Campione d'Italia 2001–2024*. Gwind S.r.l. Available at: https://www.tuttitalia.it/lombardia/21-campione-d-italia/statistiche/popolazione-andamento-demografico/ [Accessed 15.12.2025].

Hägerstrand, T. (1970). What about people in regional science. *Transport Sociology: Social aspects of transport planning*, 143–158.

Haghighi, M., & Miller, E. J. (2025). Week-long activity-based modelling: A review of the existing models and datasets and a comprehensive conceptual framework. *Transport Reviews*, *45*(1), 119–148.

Heath, E. T. (2003). Towards a model to enhance destination competitiveness: A Southern African perspective. University of Pretoria.

Horni, A., Nagel, K., & Axhausen, K. W. (2016). *Introducing matSIM. In The Multi-Agent Transport Simulation MATSim.* London: Ubiquity Press.

Horni, A., Scott, D. M., Balmer, M., & Axhausen, K. W. (2008). Location choice for shopping and leisure activities implemented in the activity-based multi-agent transport simulation matsim. *8th Swiss Transport Research Conference (STRC)*.

Ihrig, J., Jochem, P., & Szimba, E. (2024). How to apply the four-step model for 150,000 travel zones: The HIPAT model. *Transport Policy*, *155*, 150–160.

Jeewanthi, N. B., & Kumarage, A. S. (2025). Enhancing travel demand forecasting using CDR data: A stay-based integration with the four-step model. *Future Transportation*, *5*(3), 106.

Jonnalagadda, N., Freedman, J., Davidson, W. A., & Hunt, J. D. (2001). Development of microsimulation activity-based model for San Francisco: Destination and mode choice models. *Transportation Research Record: Journal of the Transportation Research Board*, *1777*(1), 25–35.

Kanasi, E., Ayilavarapu, S., & Jones, J. (2016). The aging population: demographics and the biology of aging. *Periodontology 2000*, *72*(1), 13–18.

Kerkman, K., Martens, K., & Meurs, H. (2017). A multilevel spatial interaction model of transit flows incorporating spatial and network autocorrelation. *Journal of Transport Geography*, *60*, 155–166.

Khademi, E., & Timmermans, H. (2011). Incorporating traveler response to pricing policies in comprehensive activity-based models of transport demand: Literature review and conceptualisation. *Procedia Social and Behavioral Sciences*, *20*, 594–603.

Kitamura, R. (1988). An evaluation of activity-based travel analysis. *Transportation*, *15*(1–2), 9–34.

Klinkhardt, C., Woerle, T., Briem, L., Heilig, M., Kagerbauer, M., & Vortisch, P. (2021). Using OpenStreetMap as a data source for attractiveness in travel demand models. *Transportation Research Record*, *2675*(8), 294–303.

Krešić, D., & Prebežac, D. (2011). Index of destination attractiveness as a tool for destination attractiveness assessment. *Tourism: An International Interdisciplinary Journal*, *59*(4), 497–517.

Lanzendorf, M. (2002). Mobility styles and travel behavior: Application of a lifestyle approach to leisure travel. *Transportation Research Record*, *1807*(1), 163–173.

Leong, A. M. W., Yeh, S.-S., Hsiao, Y.-C., & Huan, T. C. (2015). Nostalgia as travel motivation and its impact on tourists' loyalty. *Journal of Business Research*, *68*(1), 81–86.

Lewis, C., & Ollivaud, P. (2020). Policies for Switzerland's ageing society. Organisation for Economic Co-operation and Development (OECD). Available at: https://www.oecd.org/content/dam/oecd/en/publications/reports/2020/02/policies-for-switzerland-s-ageing-society_f05a7f36/3f8a12c6-en.pdf [Accessed 10.12.2025].

Li, T., Liao, C., Law, R., & Zhang, M. (2023). An integrated model of destination attractiveness and tourists' environmentally responsible behavior: The mediating effect of place attachment. *Behavioral Sciences*, *13*(3), 264.

Liu, S., Long, Y., Zhang, L., & Liu, H. (2021). Quantifying and characterizing urban leisure activities by merging multiple sensing big data: A case study of Nanjing, China. *Land*, *10*(11), 1214.

Liu, Y., Zhang, Y., Jin, S.-T., & Liu, Y. (2020). Spatial pattern of leisure activities among residents in Beijing, China: Exploring the impacts of urban environment. *Sustainable Cities and Society*, *52*, 101806.

Mahdi, A., Hamadneh, J., & Esztergár-Kiss, D. (2022). Modeling of travel behavior in Budapest: Leisure travelers. *Transportation Research Procedia*, *62*, 310–317.

Marrocu, E., & Paci, R. (2013). Different tourists to different destinations. Evidence from spatial interaction models. *Tourism Management*, *39*, 71–83.

Massidda, C., & Etzo, I. (2012). The determinants of Italian domestic tourism: A panel data analysis. *Tourism Management*, *33*(3), 603–610.

Mathur, A., Sherman, E., & Schiffman, L. G. (1998). Opportunities for marketing travel services to new-age elderly. *Journal of Services Marketing*, *12*(4), 265–277.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, *3*(4), 303–328.

McFadden, D. (1978). Modelling the choice of residential location. In A. Karlqvist, L. Lundqvist, F. Snickars, & J. Weibull (Eds.), *Spatial interaction theory and planning models* (pp. 75–96). North-Holland.

Menéndez, M. L., Pardo, J. A., Pardo, L., & Pardo, M. C. (1997). The jensen–shannon divergence. *Journal of the Franklin Institute*, *334*(2), 307–318.

Michigan State Highway Department. (1955). *Report on the Detroit Metropolitan Area Traffic Study*. Lansing, MI: Speaker-Hines; Thomas, State Printers.

Mitchell, R. B., & Rapkin, C. (1954). *Urban Traffic: A Function of Land Use* (p. 226). New York: Columbia University Press.

Mokras-Grabowska, J. (2018). New urban recreational spaces. attractiveness, infrastructure arrangements, identity. the example of the city of Łódź. *Miscellanea Geographica. Regional Studies on Development*, *22*(4), 219–224.

Molloy, J., & Moeckel, R. (2017). Improving destination choice modeling using location-based big data. *ISPRS International Journal of Geo-Information*, *6*(9), 291.

Moreira, P., & Iao, C. (2014). A longitudinal study on the factors of destination image, destination attraction and destination loyalty. *Journal of Social Sciences*, *3*(3), 90.

Morey, E. R., Shaw, W. D., & Rowe, R. D. (1991). A discrete-choice model of recreational participation, site choice, and activity valuation when complete trip data are not available. *Journal of Environmental Economics and Management*, *20*(2), 181–201.

Morley, C. L. (1994). Experimental destination choice analysis. *Annals of Tourism Research*, *21*(4), 780–791.

Nagel, K., Stretz, P., Pieck, M., Donnelly, R., & Barrett, C. L. (1997). *TRANSIMS traffic flow characteristics*. Available at: https://doi.org/10.48550/arXiv.adap-org/9710003 [Accessed 02.02.2026].

Notelaers, L., Verstraete, J., Vansteenwegen, P., & Tampère, C. M. J. (2024). A travel demand modeling framework based on OpenStreetMap. *Discover Civil Engineering*, *1*(1), 26.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*(5), 673–690.

Ohnmacht, T., Götz, K., & Schad, H. (2009). Leisure mobility styles in Swiss conurbations: Construction and empirical analysis. *Transportation*, *36*(2), 243–265.

Öner, Ö., & Klaesson, J. (2017). Location of leisure: The new economic geography of leisure services. *Leisure Studies*, *36*(2), 203–219.

OpenStreetMap contributors. (2026). *OpenStreetMap*. Available at: https://www.openstreetmap.org [Accessed 28.02.2026].

OpenStreetMap Wiki contributors. (2025). *Map Features*. Available at: https://wiki.openstreetmap.org/wiki/Map_features [Accessed 28.02.2026].

Overpass Turbo contributors. (2026). *Overpass Turbo - Swiss Edition*. Available at: https://overpass-turbo.osm.ch [Accessed 28.02.2026].

Pan, X., Rasouli, S., & Timmermans, H. (2021). Investigating tourist destination choice: Effect of destination image from social network members. *Tourism Management*, *83*, 104217.

Paracchini, M. L., Zulian, G., Kopperoinen, L., Maes, J., Schägner, J. P., Termansen, M., & Bidoglio, G. (2014). Mapping cultural ecosystem services: A framework to assess the potential for outdoor recreation across the EU. *Ecological Indicators*, *45*, 371–385.

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

Parmar, J., & Delle Site, P. (2025). Evaluation of approaches to internal validation of multinomial logit models: The case of personal travel mode choice. *Communications in Statistics: Case Studies, Data Analysis and Applications*, *11*(3), 316–342.

Peress, M. (2024). *Lecture Notes on Maximum Likelihood Estimation*. Department of Political Science, SUNY Stony Brook. Available at: https://www.michaelperess.com/teaching/MLE.pdf [Accessed 10.03.2026].

Petersen, F., Kuehne, H., Borgelt, C., & Deussen, O. (2022). Differentiable top-k classification learning. *International Conference on Machine Learning*, 17656–17668.

Pospíšil, J., Pospíšilová, H., & Trochtová, L. (2022). The catalogue of leisure activities: A new structured values and content based instrument for leisure research usable for social development and community planning. *Sustainability*, *14*(5), 2657.

Pozsgay, M. A., & Bhat, C. R. (2001). Destination choice modeling for home-based recreational trips: Analysis and implications for land use, transportation, and air quality planning. *Transportation Research Record*, *1777*(1), 47–54.

Pro Natura. (2023). *Pro Natura: Nature Preserves*. Pro Natura Switzerland. Available at: https://data.geo.admin.ch/browser/index.html#/collections/ch.pronatura.naturschutzgebiete/items/naturschutzgebiete?.language=de-CH&.asset=asset-naturschutzgebiete_2056-gpkg [Accessed 10.03.2026].

Rahnasto, I., & Hollestelle, M. (2024). Comparing discrete choice and machine learning models in predicting destination choice. *European Transport Research Review*, *16*(1), 43.

Ramos, C. M., & Pinto, K. (2024). How gastronomic image shapes tourism competitiveness: An approach with sentiment analysis. *Applied Sciences*, *14*(20), 9524.

Raney, B., Cetin, N., Völlmy, A., Vrtic, M., Axhausen, K. W., & Nagel, K. (2003). An agent-based microsimulation model of Swiss travel: First results. *Networks and Spatial Economics*, *3*(1), 23–41.

Recker, W. W., McNally, M. G., & Root, G. S. (1986). A model of complex travel behavior: Part I–theoretical development. *Transportation Research Part A*, *20*(4), 307–318.

Reece, W. S. (2004). Are senior leisure travelers different? *Journal of Travel Research*, *43*(1), 11–18.

Reitsamer, B. F., Brunner-Sperdin, A., & Stokburger-Sauer, N. E. (2016). Destination attractiveness and destination attachment: The mediating role of tourists' attitude. *Tourism Management Perspectives*, *19*, 93–101.

SBB. (2021). *SBB 2030 Strategy*. SBB Swiss Federal Railways (SBB). Available at: https://company.sbb.ch/en/company/about-us/strategy-2030.html [Accessed 15.03.2026].

Scherr, W., Manser, P., Joshi, C., Frischknecht, N., & Métrailler, D. (2020). Towards agent-based travel demand simulation across all mobility choices–the role of balancing preferences and constraints. *European Journal of Transport and Infrastructure Research*, *20*(4), 152–172.

Scherr, W., Bützberger, P., & Frischknecht, N. (2018). Micro Meets Macro: A Transport Model Architecture Aiming at Forecasting a Passenger Railway's Future. Swiss Federal Railways (SBB). Available at: https://www.strc.ch/2018/Scherr_EtAl.pdf [Accessed 10.03.2026].

Scherr, W., Joshi, C., Manser, P., Frischknecht, N., & Métrailler, D. (2019). An Activity-based Travel Demand Model of Switzerland Based on Choices and Constraints. Swiss Federal Railways (SBB). Available at: https://transp-or.epfl.ch/heart/2019/abstracts/hEART_2019_paper_128.pdf [Accessed 10.03.2026].

Schirpke, U., Meisch, C., Marsoner, T., & Tappeiner, U. (2018). Revealing spatial and temporal patterns of outdoor recreation in the European Alps and their surroundings. *Ecosystem Services*, *31*, 336–350.

Schlich, R., Schönfelder, S., Hanson, S., & Axhausen, K. W. (2004). Structures of leisure travel: Temporal and spatial variability. *Transport Reviews*, *24*(2), 219–237.

Scotti, F., Flori, A., Secchi, P., Arena, M., & Azzone, G. (2024). *Heterogeneous drivers of overnight and same-day visits*. Available at: https://doi.org/10.48550/arXiv.2402.05679 [Accessed 10.03.2026].

Shi, Y., Karatzoglou, A., Baltrunas, L., Larson, M., Oliver, N., & Hanjalic, A. (2012). CLiMF: Learning to maximize reciprocal rank with collaborative less-is-more filtering. *Proceedings of the 6th ACM Conference on Recommender Systems*, 139–146.

Shoemaker, S. (1989). Segmentation of the senior pleasure travel market. *Journal of Travel Research*, *27*(3), 14–21.

Simma, A., Schlich, R., & Axhausen, K. W. (2001). Destination choice modelling of leisure trips: The case of Switzerland. *Arbeitsberichte Verkehrs- und Raumplanung*, *99*.

Singh, A., & Tiwari, R. (2016). The role of destination attributes in promoting a tourist destination. *Pacific Business Review International*, *8*(10), 9–20.

Sirkis, G., Regalado-Pezúa, O., Carvache-Franco, O., & Carvache-Franco, W. (2022). The determining factors of attractiveness in urban tourism: A study in Mexico City, Buenos Aires, Bogota, and Lima. *Sustainability*, *14*(11), 6900.

Slak Valek, N., & Fotiadis, A. (2018). Is tourism really an escape from everyday life? Everyday leisure activities vs leisure travel activities of expats and Emirati nationals living in the UAE. *International Journal of Culture, Tourism and Hospitality Research*, *12*(2), 238–254.

Spinney, J. E., & Millward, H. (2013). Investigating travel thresholds for sports and recreation activities. *Environment and Planning B: Planning and Design*, *40*(3), 474–488.

State of Illinois. (1959). *Chicago Area Transportation Study: Final Report. Volume I: Survey Findings*. County of Cook, City of Chicago, IL, U.S. Department of Commerce.

Sugiyama, T., Francis, J., Middleton, N. J., Owen, N., & Giles-Corti, B. (2010). Associations between recreational walking and attractiveness, size, and proximity of neighborhood open spaces. *American Journal of Public Health*, *100*(9), 1752–1757.

Świdyńska, N., & Witkowska-Dąbrowska, M. (2021). Indicators of the tourist attractiveness of urban–rural communes and sustainability of peripheral areas. *Sustainability*, *13*(12), 6968.

Swisstopo. (2024). *Ski routes (Schneeschuhrouten)*. Swiss Federal Office of Topography. Available at: https://data.geo.admin.ch/browser/index.html#/collections/ch.swisstopo-karto.schneeschuhrouten/items/schneeschuhrouten?.language=en [Accessed 15.01.2026].

Swisstopo. (2024). *Ski routes (Skitouren)*. Swiss Federal Office of Topography. Available at: https://data.geo.admin.ch/browser/index.html#/collections/ch.swisstopo-karto.skitouren/items/skitouren?.language=en [Accessed 15.01.2026].

Swisstopo. (2025). *swissTLM3D – Topographisches Landschaftsmodell der Schweiz*. Swiss Federal Office of Topography. Available at: https://www.swisstopo.admin.ch/de/landschaftsmodell-swisstlm3d [Accessed 15.01.2026].

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288.

Töger, M., Türk, U., Östh, J., Kourtit, K., & Nijkamp, P. (2023). Inequality in leisure mobility: An analysis of activity space segregation spectra in the Stockholm conurbation. *Journal of Transport Geography*, *111*, 103638.

Train, K. E. (2009). *Discrete Choice Methods with Simulation*. Cambridge: Cambridge University Press.

United Nations. (2022). World Population Prospects 2022. Summary of Results. United Nations. Available at: https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf [Accessed 10.03.2026].

UrbiStat S.r.l. (2026). *AdminStat Germania: Karten, Analysen und Statistiken zur ansässigen Bevölkerung. Gemeinde von Büsingen am Hochrhein*. Available at: https://ugeo.urbistat.com/AdminStat/de/de/demografia/popolazione/busingen-am-hochrhein/20175339/4 [Accessed 28.01.2026].

van Zanten, B. T., Zasada, I., Koetse, M. J., Ungaro, F., Häfner, K., & Verburg, P. H. (2016). A comparative approach to assess the contribution of landscape features to aesthetic and recreational values in agricultural landscapes. *Ecosystem Services*, *17*, 87–98.

VM-UVEK. (2017). *Verkehrsmodellierung im UVEK; Daten Zustand 2017 – Verkehrszonen*. Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation (UVEK) and Bundesamt für Raumentwicklung (ARE). Available at: https://zenodo.org/records/7777651 [Accessed 23.10.2025].

Wang, K., & De Vos, J. (2025). Trends in leisure and shopping travel: From pre- to post-COVID-19 pandemic. *Travel Behaviour and Society*, *39*, 100956.

Wang, L., Shen, L., & Liu, H. (2024a). Working hours influence preferences of residents for micro features of small urban green spaces. *Scientific Reports*, *14*(1), 23948.

Wang, S., Mo, B., Zheng, Y., Hess, S., & Zhao, J. (2024b). Comparing hundreds of machine learning and discrete choice models for travel demand modeling: An empirical benchmark. *Transportation Research Part B: Methodological*, *190*, 103061.

West, R. M. (2022). Best practice in statistics: The use of log transformation. *Annals of Clinical Biochemistry*, *59*(3), 162–165.

Williams, R. (2009). Using heterogeneous choice models to compare logit and probit coefficients across groups. *Sociological Methods & Research*, *37*(4), 531–559.

Willibald, F., van Strien, M. J., Blanco, V., & Grêt-Regamey, A. (2019). Predicting outdoor recreation demand on a national scale–the case of Switzerland. *Applied Geography*, *113*, 102111.

Wilson, A. G. (1971). A family of spatial interaction models, and associated developments. *Environment and Planning A*, *3*(1), 1–32.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Wooldridge, J. M. (2016). *Introductory Econometrics: A Modern Approach*. Mason, OH: South-Western Cengage Learning.

Yates, L. A., Aandahl, Z., Richards, S. A., & Brook, B. W. (2023). Cross validation for model selection: A review with examples from ecology. *Ecological Monographs*, *93*(1), e1557.

Yoon, Y., & Uysal, M. (2005). An examination of the effects of motivation and satisfaction on destination loyalty: A structural model. *Tourism Management*, *26*(1), 45–56.

Zhang, X., Rui, J., Xia, G., Yang, J., Cai, C., & Zhao, W. (2025). Revealing disparities and driving factors in leisure activity segregation of residents and tourists: A data-driven analysis of smartphone data. *Applied Geography*, *176*, 103513.

Zhang, X., Wang, J., Kwan, M.-P., & Chai, Y. (2019). Reside nearby, behave apart? activity-space-based segregation among residents of various types of housing in Beijing, China. *Cities*, *88*, 166–180.

Zhong, M., Shan, R., Du, D., & Lu, C. (2015). A comparative analysis of traditional four-step and activity-based travel demand modeling: A case study of Tampa, Florida. *Transportation Planning and Technology*, *38*(5), 517–533.

Ziemke, D., Kaddoura, I., & Nagel, K. (2019). The MATSim open Berlin scenario: A multimodal agent-based transport simulation scenario based on synthetic demand modeling and open data. *Procedia Computer Science*, *151*, 870–877.

Ziernicka-Wojtaszek, A., & Malec, M. (2022). Evaluating local attractiveness for tourism and recreation—a case study of the communes in brzeski county, poland. *Land*, *11*(1), 39.

APPENDIX

A Exploratory data analysis of leisure trips

This appendix reports supplementary exploratory analyses of the observed leisure trips derived from the MTMC.

A.1 Trip categorisation into six leisure domains

For exploratory purposes, the extracted leisure trips are grouped into six broad activity domains: *Gastronomy*, *Visits*, *Outdoor*, *Culture*, *Sport*, and *Others*. The classification is descriptive only and does not enter the discrete choice model directly. For short trips from *WegeInland*, detailed leisure purposes are mapped to these six domains as shown in Table 15. For long trips from *Tagesreisen*, an analogous mapping is applied, while non-leisure or ambiguous purposes are excluded (Table 16). The resulting weighted domain compositions by age group and trip type are summarised in Figure 8.

Table 15: Mapping of detailed leisure-purpose categories in *WegeInland* (short trips) to broad leisure activity types used for exploratory analysis.

MTMC detailed leisure purpose (translated)	Broad activity type
Visits (relatives, acquaintances, friends)	Visits
Gastronomy visit (restaurant, bar, café, etc.)	Gastronomy
Active sport (e.g. football, jogging)	Sport
Hiking	Outdoor
Cycling	Outdoor
Non-sport outdoor activity (e.g. walking the dog)	Outdoor
Eating without a gastronomy venue (picnic, barbecue, etc.)	Outdoor
Touring/round trip (car/motorbike/train ride, etc.)	Outdoor
Cultural events and leisure facilities	Culture
Passive sport (e.g. attending a match)	Others
Medical / wellness / fitness	Others
Unpaid work	Others
Club/association activity	Others
Excursion / holiday	Others
Religion (church, cemetery, pilgrimage)	Others
Domestic leisure activities outside the home	Others
Window-shopping without purchases	Others
Other (manual entry)	Others
Multiple activities	Others
No answer / don't know	Not known

The weighted activity shares suggest that the strongest contrast emerges between *short*

Table 16: Mapping of general day-trip purposes in *Tagesreisen* (long/day trips) to broad leisure activity types used for exploratory analysis.

MTMC detailed day-trip purpose (translated)	Broad activity type
Visits (relatives, acquaintances)	Visits
Gastronomy visit	Gastronomy
Active sport	Sport
Hiking	Outdoor
Cycling	Outdoor
Non-sport outdoor activities	Outdoor
Touring/round trip	Outdoor
Cultural event / leisure facilities	Culture
Passive sport	Others
Excursion / holiday	Others
Religion (church, cemetery, pilgrimage)	Others
Trip not selected / no answer / don't know	Excluded
Education trip / camp	Excluded
Shopping	Excluded
Medical treatment	Excluded
Business trip	Excluded
Accompanying a business trip	Excluded
Accompanying a private trip	Excluded
Other	Excluded

and *long* leisure trips. Across both age groups, the share of *Visits* increases for long trips, rising from around one quarter of short trips to roughly one-third or more of long trips. *Gastronomy*, by contrast, is much more prominent among short trips and falls sharply for long/day trips, where it remains below 5% in both age groups. The share of *Culture*, *Outdoor* and *Others* is also somewhat higher for long trips. *Sport* tends to decline for long trips. Differences between age groups are also visible, although they are generally less pronounced than the contrast between short and long trips. In both trip types, older adults show a higher share of *Outdoor* activities and a lower share of *Sport* than younger individuals. Older adults also show a somewhat higher share of *Others*. *Gastronomy* is slightly more prominent among older adults, while *Visits* are slightly more frequent among younger individuals.

A.2 Spatial distribution of leisure destinations by segment

This subsection visualises the spatial distribution of observed leisure destinations across the four segments (young/old × short/long). All maps are based on `WP`-weighted destination counts aggregated to NPVM traffic zones. To facilitate visual comparison across segments, the weighted destination-count surfaces are rescaled within each segment to the [0,1] interval (Figure 9). Difference maps then display contrasts between these segment-

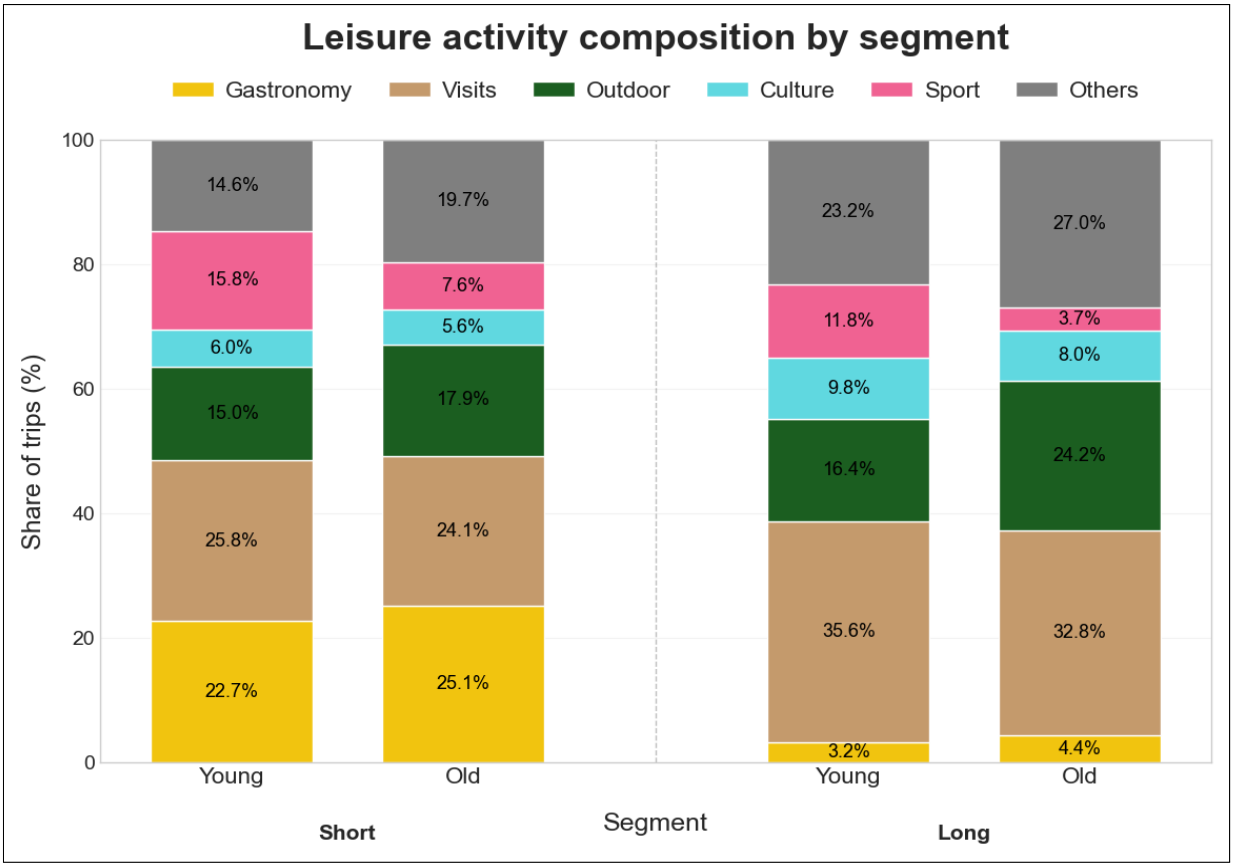


Figure 8: Weighted composition of broad leisure activity domains by segment.

specific normalised surfaces (Figure 10). In addition, Table 17 reports Pearson and Spearman correlations of the raw `WP`-weighted destination counts across zones for the main pairwise segment comparisons.

Table 17: Pearson and Spearman correlations of destination counts across zones.

Comparison (across zones)	Pearson	Spearman
Young vs. Old | Short trips	0.5692	0.4384
Young vs. Old | Long trips	0.9100	0.5043
Short vs. Long | Young group	0.2853	0.1424
Short vs. Long | Old group	0.2216	0.1584

A first descriptive point concerns the uneven spatial coverage of observed destinations across segments. Out of the 7,978 NPVM traffic zones, 169 do not contain any observed leisure destination in any of the four segments. The number of zero-destination zones varies substantially by segment, from 713 for Young–Short and 2,856 for Old–Short to 6,012 for Young–Long and 7,044 for Old–Long. Long/day-trip leisure is more spatially concentrated than short leisure mobility. In both age groups, the spatial association between short and long trips is weak, indicating that day-to-day leisure destinations and longer day-trip destinations follow markedly different geographies. Visually, long-trip

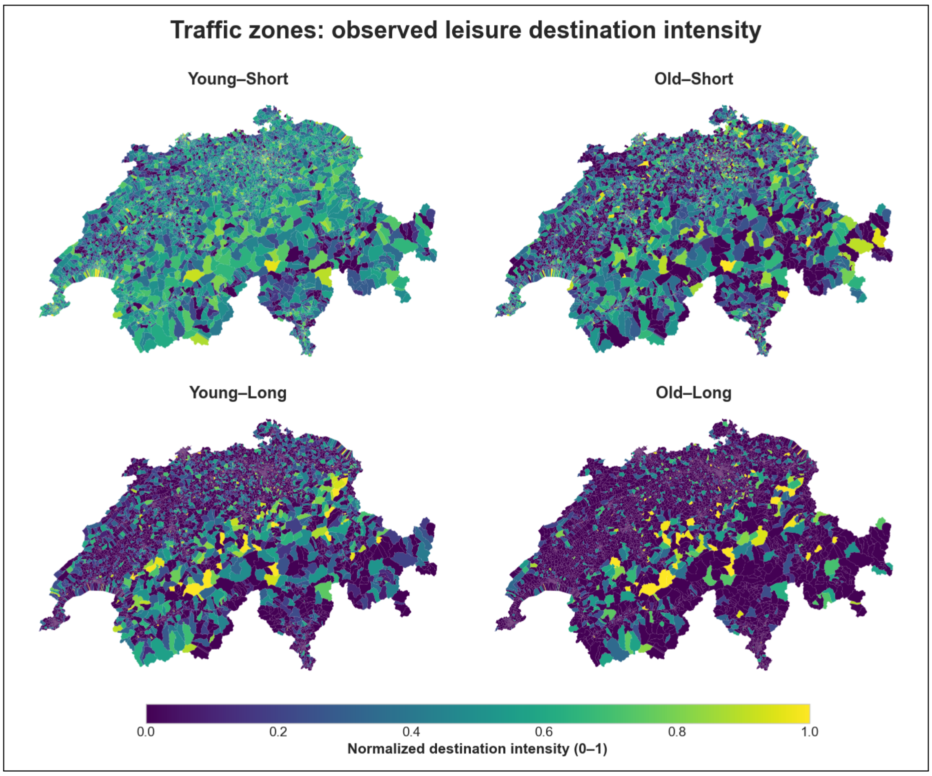


Figure 9: NPVM traffic zones: normalised intensity of observed leisure destinations for the four segments (Young–Short, Old–Short, Young–Long, Old–Long), based on `WP`-weighted destination counts.

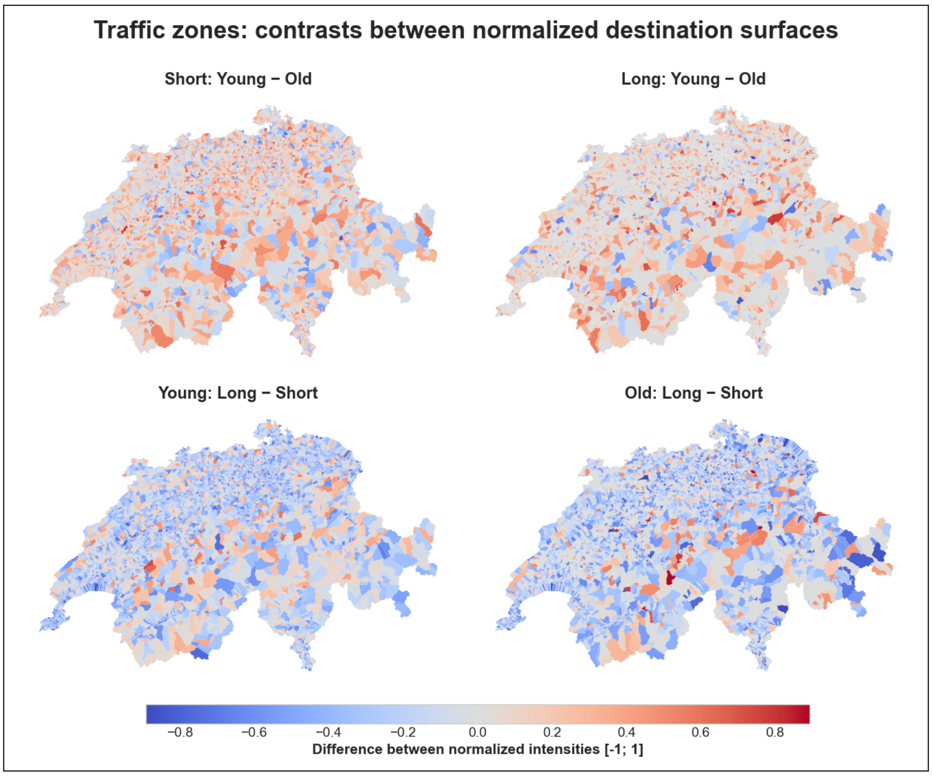


Figure 10: NPVM traffic zones: pairwise contrasts between segment-specific normalised destination-intensity surfaces (young–old and long–short). Positive and negative values indicate the relative dominance of one segment over the other.

destinations appear more concentrated in a limited number of zones, with particularly strong intensity along the Alpine chain, especially in Central Switzerland. Differences between younger and older individuals appear less pronounced than the contrast between short and long trips. For short trips the correlation is lower and the maps suggest that destinations of older adults are somewhat more localised, whereas the Young–Short surface appears more spatially diffuse. For long trips, the difference between age groups appears smaller overall. At the same time, the young–old difference maps still show several zones with positive or negative values, suggesting that age-related differences in destination patterns are present.

A.3 Spatial clustering of leisure destinations

To describe local clustering in observed leisure destinations, I compute Local Moran's I, one of the Local Indicators of Spatial Association (LISA) introduced by Anselin (1995), on log-transformed `WP`-weighted destination counts using queen contiguity neighbourhoods. Rather than the four full segments, the analysis is shown from four aggregated perspectives: *Young* (short+long), *Old* (short+long), *Short* (young+old), and *Long* (young+old). The resulting Local Moran cluster map is shown in Figure 11. Each zone is classified into one of five categories. *HH hotspots* identify zones with high observed destination intensity surrounded by neighbouring zones that are likewise high. *LL coldspots* indicate the opposite configuration, that is, low-value zones surrounded by similarly low-value neighbours. *HL* and *LH* denote spatial outliers: in the first case, a high-value zone surrounded by low-value neighbours; in the second, a low-value zone surrounded by high-value neighbours. Grey areas are not statistically significant at the 5% level. Significant clusters account for 20.17% of zones in the *Young* view, 17.11% in the *Old* view, 21.46% in the *Short* view, and 12.17% in the *Long* view. Visually, the most distinctive pattern emerges in the *Long* perspective, where HH hotspots appear more clearly concentrated along the Alpine chain. By contrast, differences among the other three perspectives are less pronounced. In particular, the *Young* and *Old* maps do show some differences, but the main hotspot and coldspot areas remain broadly similar or spatially close to one another. Overall, this descriptive evidence suggests that some areas consistently exhibit higher observed leisure-destination intensity than others across aggregation perspectives.

A.4 Spatial distribution of leisure destinations by leisure type

This subsection maps observed destinations separately by the broad leisure categories. For each category, `WP`-weighted destination counts are aggregated to NPVM traffic zones and displayed as robustly normalised intensity surfaces. Separate panels are shown for short trips (*WegeInland*) and long/day trips (*Tagesreisen*). Together, Figures 12 and 13 provide a descriptive view of how the geography of observed leisure destinations varies.

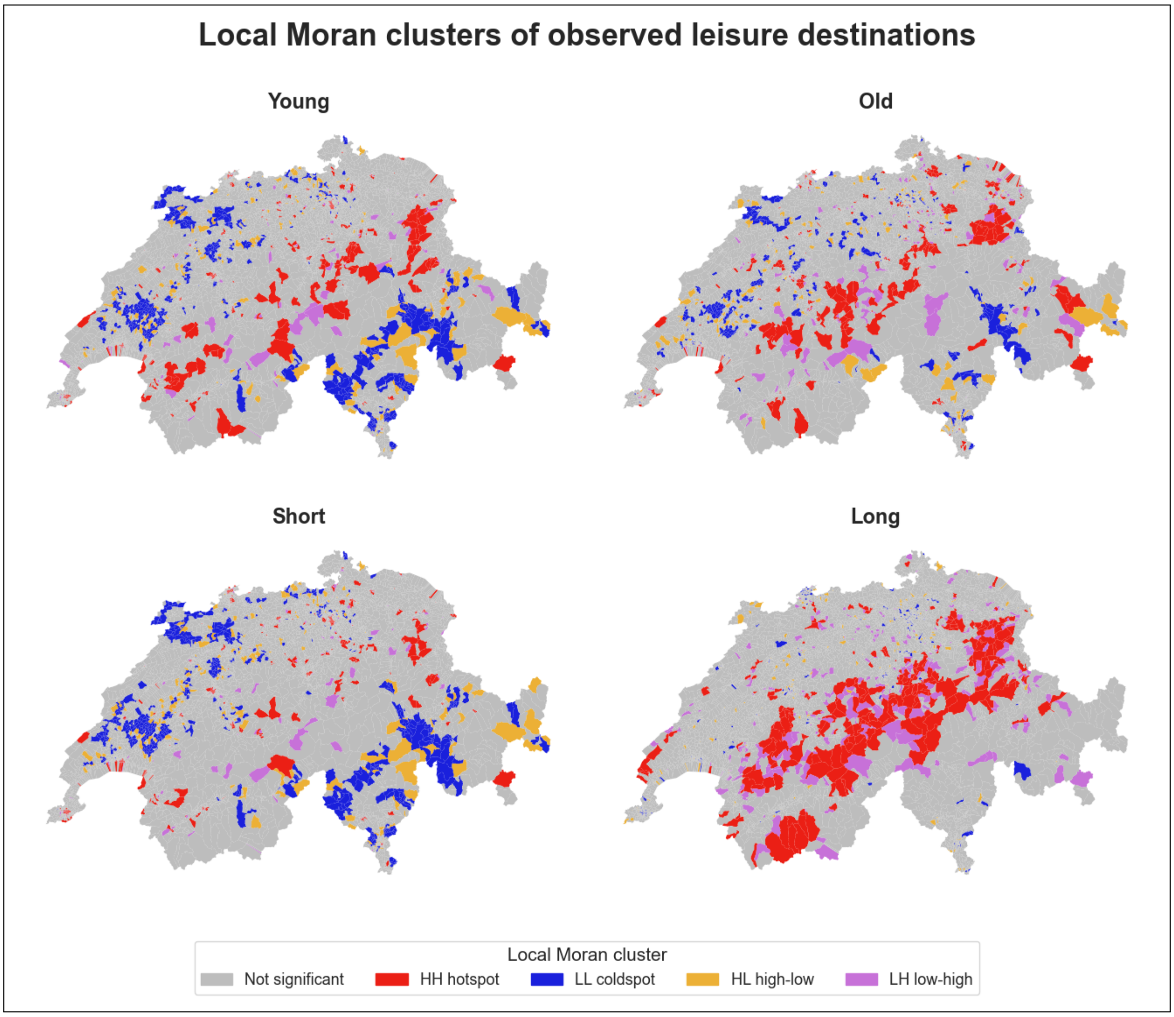


Figure 11: Local Moran cluster maps of WP-weighted observed leisure destinations shown from four aggregated perspectives: Young, Old, Short, and Long.

For both short and long trips, the maps suggest some differences in the spatial distribution of destinations across leisure categories. Visits appear comparatively widespread across the country in both cases. For short trips, *Outdoor* and *Others* also appear relatively widely distributed, whereas *Gastronomy*, *Sport*, and *Culture* seem more spatially concentrated. For long/day trips, *Sport*, *Outdoor*, and *Others* appear more concentrated in Alpine areas, while *Visits* seems more spatially diffuse in the short-trip case. Overall, these descriptive maps further support the idea that leisure mobility is heterogeneous not only across age groups and trip types, but also across broad activity domains.

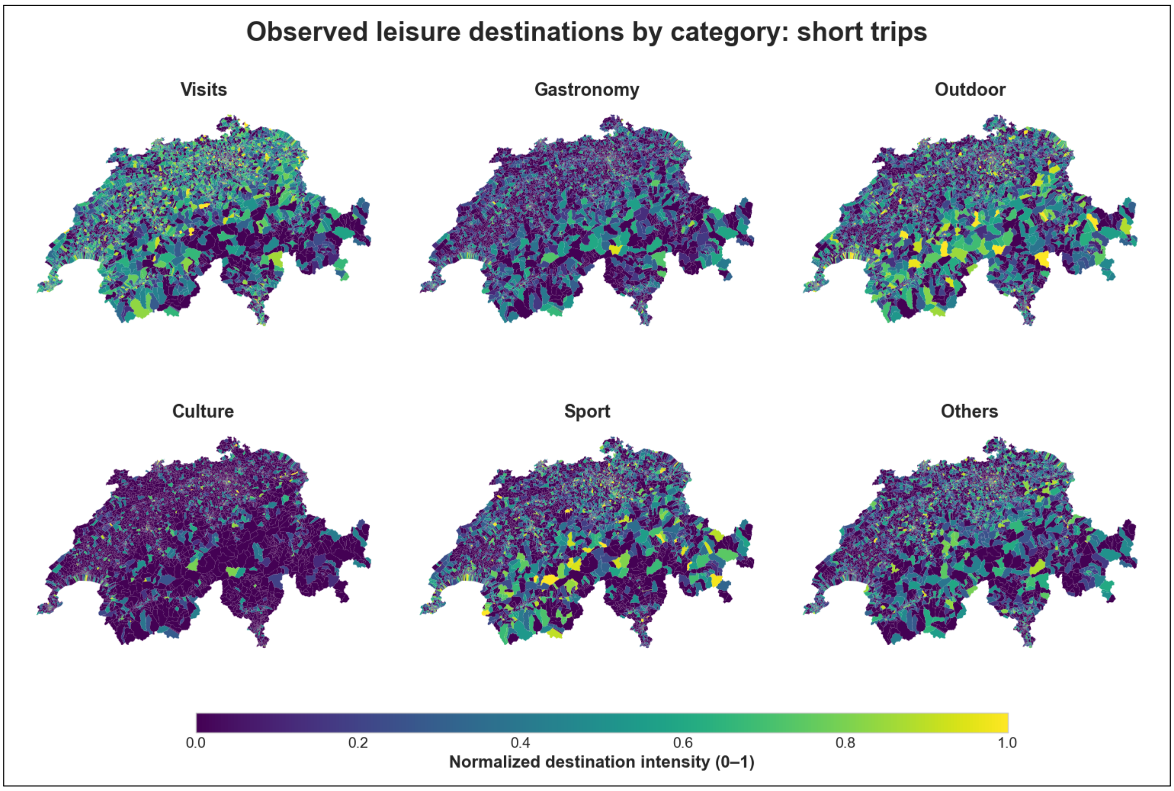


Figure 12: Category-specific maps of weighted leisure destinations for short trips.

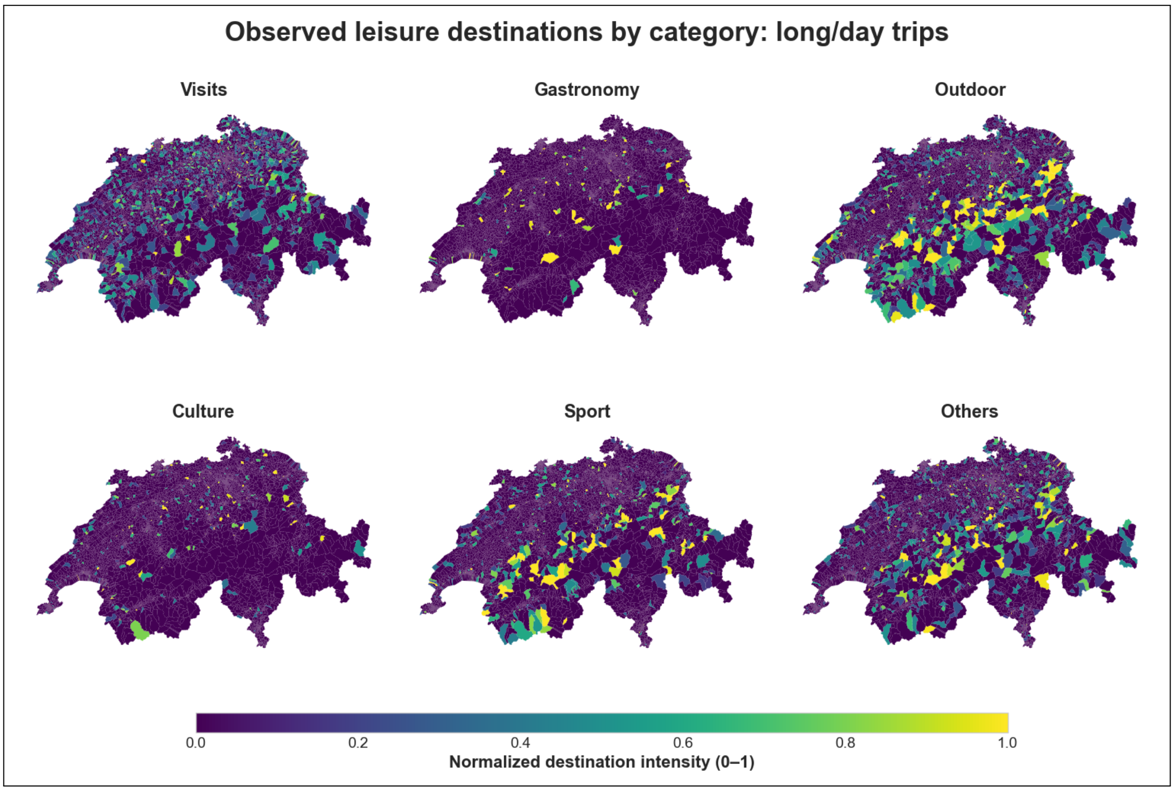


Figure 13: Category-specific maps of weighted leisure destinations for long/day trips.

B Overpass Turbo queries (OSM)

This appendix documents the Overpass Turbo queries used to extract OpenStreetMap (OSM) features for the construction of the supply-side variables. The objects downloaded from Overpass Turbo are converted from GeoJSON to GeoPackage format and reprojected to the Swiss reference system (EPSG:2056) to ensure consistency with all other spatial datasets used in the analysis. All queries follow the same general logic. First, a common *prefix* defines the extraction area, which covers Switzerland, the Principality of Liechtenstein, and the two foreign enclaves Büsingen am Hochrhein and Campione d'Italia. Second, each feature is defined through a category-specific tag selection block. Third, a common *suffix* specifies the output mode.

The common area-definition prefix used throughout the queries is:

```
[out:json][timeout:180];

/* Areas: Switzerland + Liechtenstein + two enclaves */
area["ISO3166-1"="CH"]->.ch;
area["ISO3166-1"="LI"]->.li;
area["name"="Campione d'Italia"]["boundary"="administrative"]->.ca;
area["name"="Büsingen am
↪  Hochrhein"]["boundary"="administrative"]->.bu;
(.ch;.li;.ca;.bu;)->.a;
```

The category-specific query body is then inserted between this prefix and one of the following suffixes. In most cases, full geometries were retained, using:

```
out body;
>;
out skel qt;
```

Where point representations were sufficient (e.g. for point-based zonal aggregation), the alternative suffix was:

```
out center;
```

In practice, each complete Overpass query is obtained by combining the common prefix, one of the category-specific bodies reported below, and the relevant suffix.

F1 Gastronomy

```
nwr["amenity"~"^(restaurant|fast_food|cafe|pub|bar|ice_cream|food_c
  ↪  ourt|biergarten)$"](area.a);
```

F3 Outdoor recreation: Rivers

```
(
  way["waterway"~"^(river|stream|canal)$"](area.a);
  relation["waterway"~"^(river|stream|canal)$"](area.a);
);
```

F3 Outdoor recreation: Hard outdoor features

```
(
  nwr["aerialway"~"^(zip_line|gondola|cable_car)$"](area.a);
  nwr["tourism"~"^(alpine_hut|viewpoint|wilderness_hut)$"](area.a);
  nwr["natural"~"^(glacier|cave_entrance)$"](area.a);
  nwr["waterway"="waterfall"](area.a);
);
```

F3 Outdoor recreation: Soft outdoor features

```
(
  nwr["leisure"~"^(beach_resort|bird_hide|dog_park|firepit|park|pic
    ↪  nic_table|playground|marina)$"](area.a);
  nwr["tourism"="picnic_site"](area.a);
);
```

F4 Cultural opportunities

```
(
  nwr["tourism"~"^(aquarium|gallery|museum|zoo)$"](area.a);
  nwr["amenity"~"^(arts_centre|cinema|conference_centre|events_venu
    ↪  e|exhibition_centre|museum|music_venue|planetarium|theatre|li
    ↪  brary)$"](area.a);
);
```

F5 Sport opportunities

```
(
  nwr["building"~"^(pavilion|riding_hall|sports_hall|sports_centre|
  ↪  stadium)$"](area.a);

  nwr["leisure"~"^(dance|disc_golf_course|fitness_centre|fitness_st
  ↪  ation|horse_riding|ice_rink|miniature_golf|paddling_pool|pitc
  ↪  h|slipway|sports_centre|swimming_area|track)$"](area.a);
);
```

F5 Sport opportunities: Linear infrastructure (via ferrata)

```
(
  way["highway"="via_ferrata"](area.a);
  relation["highway"="via_ferrata"](area.a);
);
```

F6 Other facilities

```
(
  nwr["amenity"~"^(brothel|casino|gaming|love_hotel|nightclub|strip
  ↪  club|swingerclub|community_centre|social_centre|grave_yard|mo
  ↪  nastery|place_of_worship|kneipp_water_cure)$"](area.a);

  nwr["leisure"~"^(adult_gaming_centre|amusement_arcade|escape_game
  ↪  |water_park|hackerspace)$"](area.a);

  nwr["tourism"="theme_park"](area.a);

  nwr["shop"~"^(massage)$"](area.a);
);
```

F10 Supporting infrastructure and services

```
(
  nwr["amenity"~"^(fuel|charging_station|atm|toilets|drinking_water
  ↪  |fountain|water_point|watering_place)$"](area.a);
);
```

C Summary of OSM-derived amenities and counts

Table 18 summarises the main OpenStreetMap feature groupings used in the construction of the supply-side features. Depending on the feature, entries are reported either as raw OSM `key=value` pairs or as grouped categories used in the empirical workflow. Totals refer to the final retained OSM datasets after preprocessing and deduplication.

Table 18: Summary of retained OSM amenity groupings and counts by attractivity feature.

Key / Type	**Value / Category**	**Count**
F1 Gastronomy (total: 28,882)		
amenity	{restaurant, food_court}	18,466
amenity	cafe	3,694
amenity	{bar, pub, biergarten}	3,366
amenity	fast_food	3,056
amenity	ice_cream	300
F3 Outdoor recreation – hard features (total: 8,648)		
aerialway	{cable_car, gondola, zip_line}	505
tourism	viewpoint	4,289
tourism	{alpine_hut, wilderness_hut}	874
natural	glacier	935
natural	cave_entrance	592
waterway	waterfall	1,453
F3 Outdoor recreation – soft features (total: 35,909)		
leisure	playground	13,794
leisure, tourism	leisure = picnic_table; tourism = picnic_site	13,453
leisure	park	4,213
leisure	firepit	3,954
leisure	marina	451
leisure	dog_park	132
leisure	bird_hide	106
leisure	beach_resort	45
F4 Cultural opportunities (total: 4,245)		
tourism	museum	1,257
amenity	library	1,205
amenity	theatre	474
tourism	gallery	327
amenity	cinema	268
amenity	events_venue	243
amenity	arts_centre	224
tourism	zoo	161

Continued on next page

Key / Type	Value / Category	Count
amenity	conference_centre	44
amenity	exhibition_centre	19
amenity	music_venue	15
amenity	planetarium	6
tourism	aquarium	2
F5 Sport opportunities (total: 40,111)		
leisure	pitch	28,082
leisure	{fitness_centre, fitness_station}	4,852
building	sports_centre	3,494
leisure	track	1,888
leisure	slipway	428
leisure	horse_riding	273
leisure	miniature_golf	232
leisure	swimming_area	192
building	sports_hall	171
leisure	ice_rink	156
building	pavilion	151
leisure	dance	113
building	stadium	57
building	riding_hall	43
leisure	disc_golf_course	15
F6 Other facilities (total: 10,219)		
multiple tags	Nightlife: amenity = {brothel, casino, love_hotel, nightclub, stripclub, swingerclub}	505
multiple tags	Social: amenity = {community_centre, social_centre}	1,027
multiple tags	Spiritual: amenity = {grave_yard, monastery, place_of_worship}	7,937
multiple tags	Fun: leisure = {adult_gaming_centre, amusement_arcade, escape_game, water_park, hackerspace}; tourism = theme_park; amenity = gaming	297
multiple tags	Wellness: amenity = kneipp_water_cure; shop = massage	453
F10 Supporting infrastructure (total: 40,053)		
amenity	fuel	3,780
amenity	charging_station	3,396
amenity	atm	2,385
amenity	toilets	6,447
amenity	drinking_water	14,429
amenity	fountain	8,031
amenity	water_point	164
amenity	watering_place	1,421

D Plausibility checks for OSM-derived indicators

This appendix summarises the plausibility checks conducted for the OSM-derived indicators used in the construction of the destination-side feature blocks. Two main types of official comparison are employed. First, for *count-based* OSM indicators, the comparison is carried out against official counts of establishments, facilities, or objects also aggregated at the zonal level. In the case of economic activities, this corresponds to the number of workplaces (*Arbeitsstätten*) from STATENT. Second, for *area-based* OSM proxies, the comparison is carried out against full-time equivalent workers (FTE/VZÄ) from STATENT, because the purpose is to approximate the size or capacity of a sector. This is justified by the fact that both surface area and FTE can be interpreted, albeit imperfectly, as proxies for the overall scale of the corresponding activity. Sectoral comparisons based on STATENT rely on the Swiss NOGA classification (BFS, 2008). Where no suitable employment-based benchmark exists, OSM-derived counts are compared with spatial objects extracted from official topographic datasets such as swissTLM3D or GLAMOS.

Table 19 reports the resulting Pearson and Spearman correlations across Swiss traffic zones.

Table 19: Summary of plausibility checks for selected OSM-derived indicators against official benchmark datasets.

OSM-derived variable	**Official comparison**	**Pearson**	**Spearman**
`F1_gastr_count`	STATENT workplaces, gastronomy sector (NOGA code: 56)	0.689	0.653
`F1_gastr_area`	STATENT full-time equivalents, gastronomy sector (NOGA code: 56)	0.603	0.621
`F3_outdoor_hard_count` – glaciers	GLAMOS glacier inventory	0.955	0.939
`F3_outdoor_hard_count` – aerialways	swissTLM3D cableways	0.876	0.858
`F3_outdoor_hard_count` – waterfalls	swissTLM3D single objects – waterfalls	0.665	0.523
`F3_outdoor_hard_count` – caves	swissTLM3D single objects - caves	0.400	0.439
`F3_outdoor_hard_count` – huts	swissTLM3D single objects – huts / shelters	0.441	0.365
OSM `leisure=park` count	swissTLM3D public park areas	0.428	0.424
`F4_cult_count`	STATENT workplaces, cultural sectors (NOGA codes: 90 + 91)	0.326	0.267
`F4_cult_area_total`	STATENT full-time equivalents, cultural sectors (NOGA codes: 90 + 91)	0.309	0.285
`F5_sport_count`	swissTLM3D sport facilities and leisure areas	0.764	0.739
`F5_sport_area_total`	STATENT full-time equivalents, sport and recreation sector (NOGA code: 93)	0.293	0.250

E Variable inventory (traffic zones)

Table 20 reports the complete inventory of constructed analysis variables stored in the NPVM traffic-zone layer (`TZ.gpkg`). The prefix $F1$–$F10$ indicates the main feature block to which a variable is assigned. The label `off` denotes an official-data counterpart used alongside, or for comparison with, an OSM-derived indicator. The suffix `dens` indicates that the variable is normalised by zone area, while `div` or `diversity` refers to a diversity measure. The suffix `norm` denotes variables rescaled to the $[0, 1]$ interval. Finally, `sum` and `mean` indicate simple aggregate summaries of the corresponding components. The sequence of elements in the variable names reflects the order of the operations applied during variable construction.

Table 20: Complete inventory of constructed variables stored in `TZ.gpkg`.

No.	Variable	No.	Variable
1	F1_gastr_FTE	68	F4_cult_music_venue_count
2	F1_gastr_area_total	69	F4_cult_off_count
3	F1_gastr_count	70	F4_cult_planetarium_count
4	F1_gastr_div	71	F4_cult_theatre_count
5	F1_gastr_div_norm	72	F4_cult_zoo_count
6	F1_gastr_off_counts	73	F5_sport_FTE
7	F2_hh_total	74	F5_sport_area_total
8	F2_pop_total	75	F5_sport_count
9	F3_cable_off_count	76	F5_sport_div
10	F3_cave_off_count	77	F5_sport_div_norm
11	F3_glacier_off_count	78	F5_sport_offTLM_count
12	F3_hut_off_count	79	F5_sport_off_count
13	F3_lake_area	80	F5_sport_out_div
14	F3_outdoor_LUmix	81	F5_sport_out_div_norm
15	F3_outdoor_aerialway_count	82	F5_sport_out_length
16	F3_outdoor_alpine_hut_count	83	F6_others_count
17	F3_outdoor_cablecar_count	84	F6_others_diversity
18	F3_outdoor_cave_count	85	F6_others_diversity_norm
19	F3_outdoor_glacier_count	86	F6_others_fun
20	F3_outdoor_gondola_count	87	F6_others_nightlife
21	F3_outdoor_hard_count	88	F6_others_social
22	F3_outdoor_hard_count_off	89	F6_others_spiritual
23	F3_outdoor_hard_diversity	90	F6_others_wellness
24	F3_outdoor_hard_diversity_norm	91	F7_div_sum
25	F3_outdoor_hut_count	92	F7_div_sum_norm
26	F3_outdoor_lake_bin	93	F7_norm_mean
27	F3_outdoor_lake_shore	94	F7_norm_sum
28	F3_outdoor_protected	95	F7_norm_sum_norm
29	F3_outdoor_protected_area	96	F7_outdoor_mean
30	F3_outdoor_river	97	F7_outdoor_mean_norm
31	F3_outdoor_soft_area_imp	98	F7_outdoor_sum
32	F3_outdoor_soft_area_poly	99	F7_outdoor_sum_norm
33	F3_outdoor_soft_beach_resort_count	100	F7_urban_mean

No.	Variable	No.	Variable
34	F3_outdoor_soft_bird_hide_count	101	F7_urban_mean_norm
35	F3_outdoor_soft_count	102	F7_urban_sum
36	F3_outdoor_soft_diversity	103	F7_urban_sum_norm
37	F3_outdoor_soft_diversity_norm	104	F8_POI_natural_dens
38	F3_outdoor_soft_dog_park_count	105	F8_POI_total_dens
39	F3_outdoor_soft_firepit_count	106	F8_POI_urban_dens
40	F3_outdoor_soft_leisureuse_off_count	107	F8_cult_area_total_dens
41	F3_outdoor_soft_marina_count	108	F8_cult_count_dens
42	F3_outdoor_soft_off_count	109	F8_gastr_area_total_dens
43	F3_outdoor_soft_park_count	110	F8_gastr_count_dens
44	F3_outdoor_soft_picnic_site_count	111	F8_hh_total_dens
45	F3_outdoor_soft_picnic_table_count	112	F8_lake_area_dens
46	F3_outdoor_soft_playground_count	113	F8_lengths_dens
47	F3_outdoor_soft_public_parks_off_count	114	F8_others_count_dens
48	F3_outdoor_viewpoint_count	115	F8_others_fun_dens
49	F3_outdoor_walk	116	F8_others_nightlife_dens
50	F3_outdoor_waterfall_count	117	F8_others_social_dens
51	F3_outdoor_wilderness_hut_count	118	F8_others_spiritual_dens
52	F3_outdoor_zipline_count	119	F8_others_wellness_dens
53	F3_waterfall_off_count	120	F8_outdoor_hard_count_dens
54	F4_cult_FTE	121	F8_outdoor_lake_shore_dens
55	F4_cult_aquarium_count	122	F8_outdoor_protected_area_dens
56	F4_cult_area_total	123	F8_outdoor_river_dens
57	F4_cult_arts_centre_count	124	F8_outdoor_soft_area_imp_dens
58	F4_cult_cinema_count	125	F8_outdoor_soft_area_poly_dens
59	F4_cult_conference_centre_count	126	F8_outdoor_soft_count_dens
60	F4_cult_count	127	F8_outdoor_walk_dens
61	F4_cult_diversity	128	F8_pop_total_dens
62	F4_cult_diversity_norm	129	F8_sport_area_total_dens
63	F4_cult_events_venue_count	130	F8_sport_count_dens
64	F4_cult_exhibition_centre_count	131	F8_sport_out_length_dens
65	F4_cult_gallery_count	132	F8_superinfra_dens
66	F4_cult_library_count	133	F10_superinfra
67	F4_cult_museum_count		

F Utility-based accessibility specification

This appendix reports the mode-specific utility specifications used to construct the accessibility measure adopted in this thesis. These specifications are inherited from the operational SIMBA MOBi framework and are therefore not estimated or recalibrated here. For each origin–destination pair (i, j) and each mode $m \in \{\text{walk}, \text{bike}, \text{car}, \text{PT}\}$, a systematic travel utility U_{ijm} is computed from OD skim matrices and zonal accessibility attributes. Distances are measured in kilometres and travel times in minutes. The variables entering the specification are summarised in Table 21.

Table 21: Variables used in the utility-based accessibility specification.

Variable	Description
dist_{ij}	Car-network distance between origin i and destination j (km)
TT_{ij}^{car}	Car travel time between i and j (min)
AC_i^{car}, AC_j^{car}	Car access time at origin and destination zone (min)
PC_j^{car}	Parking cost at destination zone j (CHF/h)
TT_{ij}^{pt}	Total public-transport travel time (min)
TT_{ij}^{train}	Train component of PT travel time (min)
TT_{ij}^{bus}	Non-train component of PT travel time (min)
ACC_{ij}^{pt}	PT access time (min)
EGR_{ij}^{pt}	PT egress time (min)
tr_{ij}	Number of PT transfers
freq_{ij}	PT service frequency (departures per hour)

For car and public transport, distance enters the specification through a four-band piecewise-linear formulation with breakpoints at 15, 50, and 100 km. In practice, the total car-network distance dist_{ij} is split into four incremental components: the portion between 0 and 15 km, 15 and 50 km, 50 and 100 km, and above 100 km, denoted respectively by $\text{dist}_{ij}^{0\text{–}15}$, $\text{dist}_{ij}^{15\text{–}50}$, $\text{dist}_{ij}^{50\text{–}100}$, and dist_{ij}^{100+}. For example, if $\text{dist}_{ij} = 70$ km, then the four components are 15, 35, 20, and 0 km. The mode-specific systematic utilities are:

$$U_{ij}^{bike} = -0.25 - 0.150 \cdot \frac{\text{dist}_{ij}}{0.21667}, \tag{26}$$

$$U_{ij}^{walk} = +2.30 - 0.100 \cdot \frac{\text{dist}_{ij}}{0.078336}, \tag{27}$$

$$\begin{aligned} U_{ij}^{car} = {} & -0.40 - 0.053 \cdot \text{TT}_{ij}^{car} - 0.040 \cdot \text{dist}_{ij}^{0\text{–}15} - 0.040 \cdot \text{dist}_{ij}^{15\text{–}50} + 0.015 \cdot \text{dist}_{ij}^{50\text{–}100} \\ & + 0.010 \cdot \text{dist}_{ij}^{100+} - 0.047 \cdot (\text{AC}_i^{car} + \text{AC}_j^{car}) - 0.135 \cdot (2 \cdot \text{PC}_j^{car}), \end{aligned} \tag{28}$$

$$\begin{aligned} U_{ij}^{pt} = {} & +0.75 - 0.042 \cdot \text{TT}_{ij}^{bus} - 0.0378 \cdot \text{TT}_{ij}^{train} - 0.015 \cdot \text{dist}_{ij}^{0\text{–}15} \\ & - 0.015 \cdot \text{dist}_{ij}^{15\text{–}50} + 0.005 \cdot \text{dist}_{ij}^{50\text{–}100} + 0.025 \cdot \text{dist}_{ij}^{100+} \\ & - 0.050 \cdot (\text{ACC}_{ij}^{pt} + \text{EGR}_{ij}^{pt}) - 0.014 \cdot \left(\frac{60}{\text{freq}_{ij}}\right) - 0.227 \cdot \text{tr}_{ij}. \end{aligned} \tag{29}$$

G Spatial distribution of zonal features

This appendix reports the spatial distribution of the 13 destination-side indicators retained in the final attractivity specification. All maps are shown after the $\log(1+x)$ preprocessing transformation, except for the land-use mix indicator.

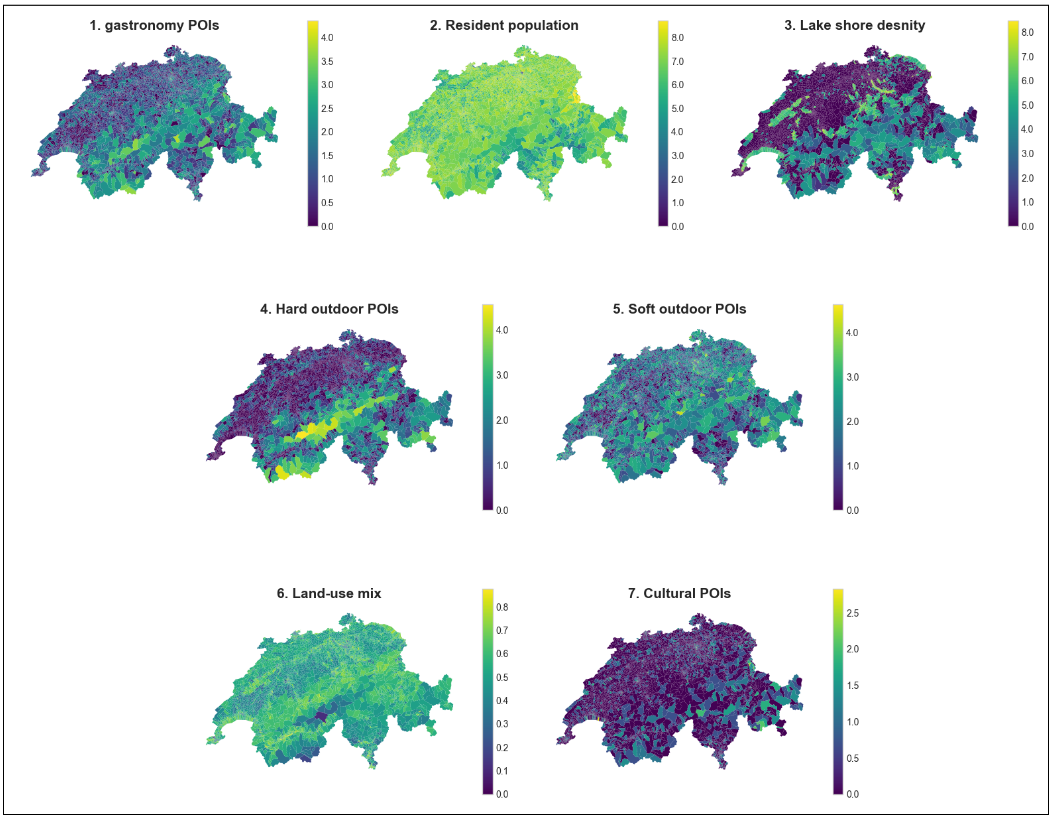


Figure 14: Spatial distribution of selected zonal features (part I).

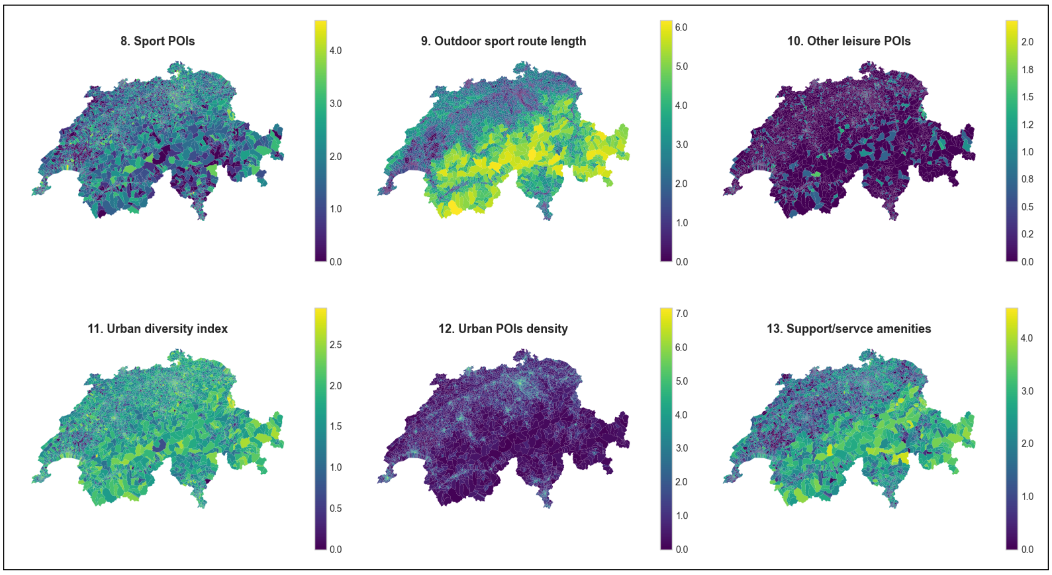


Figure 15: Spatial distribution of selected zonal features (part II).

H Choice-set sampling diagnostics

This appendix reports two diagnostics that motivate the choice-set sampling design adopted in the main analysis. First, Figure 16 shows the distribution of chosen-destination accessibility (EMU_{ij}) by segment, comparing the training and out-of-sample subsamples. The figure highlights a clear difference between short- and long-trip segments, whereas differences across age groups appear less pronounced. Second, Figure 17 reports the sensitivity analysis used to select the number of sampled non-chosen alternatives K. Predictive performance is evaluated on the inner validation sample using the full destination set at evaluation time. Across segments, performance improves with larger sampled choice sets but tends to stabilise around $K = 1000$.

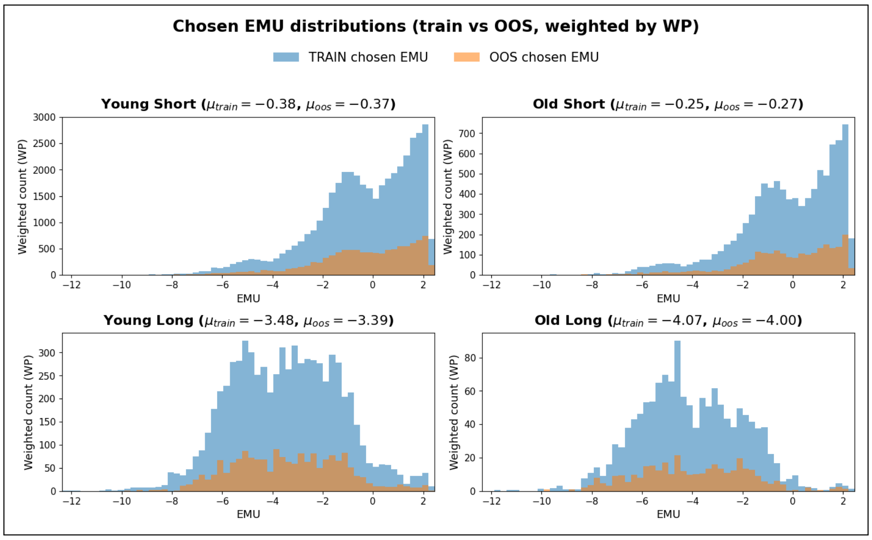


Figure 16: Chosen-destination accessibility (EMU_{ij}) distributions by segment.

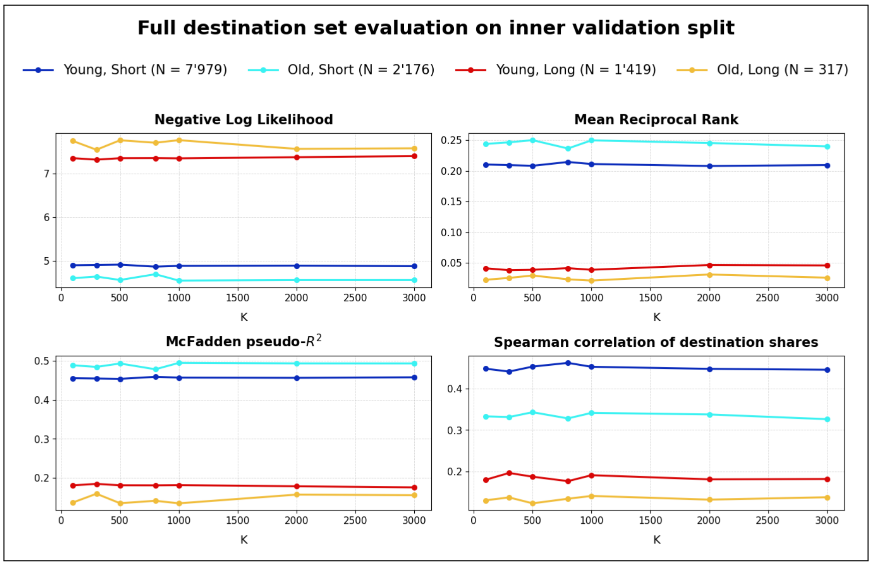


Figure 17: Sensitivity of validation performance to K.

I Coefficient significance estimation

This appendix describes how approximate coefficient uncertainty and pairwise coefficient comparisons are obtained after model estimation. Its purpose is to clarify how approximate standard errors, z-statistics, p-values, and coefficient-difference tests are derived from the estimated models. The procedure follows the logic of post-estimation Hessian-based inference and Wald-type coefficient contrasts (Peress, 2024; Wooldridge, 2010). These inferential quantities are used in this thesis as approximate post-estimation diagnostics supporting coefficient interpretation, rather than as exact finite-sample inference.

Hessian-based approximation. Let $\hat{\theta} = (\hat{\alpha}_{\mathrm{EMU}}, \hat{\beta}_1, \ldots, \hat{\beta}_H)'$ denote the vector of estimated parameters for one segment-specific model. Approximate coefficient uncertainty is derived from the local curvature of the weighted objective function around the estimated optimum. If $\mathcal{L}(\theta)$ denotes the weighted log-likelihood, the observed information matrix is defined as

$$\mathcal{I}(\hat{\theta}) = -\left.\frac{\partial^2 \mathcal{L}(\theta)}{\partial\theta\,\partial\theta'}\right|_{\theta=\hat{\theta}}, \qquad \widehat{\mathrm{Var}}(\hat{\theta}) \approx \mathcal{I}(\hat{\theta})^{-1}, \qquad \mathrm{se}(\hat{\theta}_h) = \sqrt{\widehat{\mathrm{Var}}(\hat{\theta}_h)}. \tag{30}$$

Equivalently, since the implementation differentiates the weighted negative log-likelihood, the observed information is obtained as the Hessian of that objective evaluated at the estimated optimum. In practice, the implementation applies a very small ridge adjustment for numerical stability and then uses a symmetric pseudo-inverse to obtain an approximate variance-covariance matrix. Approximate significance is assessed for $H_0 : \theta_h = 0$ through

$$z_h = \frac{\hat{\theta}_h}{\mathrm{se}(\hat{\theta}_h)}, \qquad p_h = 2\big[1 - \Phi(|z_h|)\big], \tag{31}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

Pairwise coefficient comparisons within segments. To assess whether two coefficients within the same segment differ from one another, I consider the null hypothesis that their difference is zero, $H_0 : \beta_a - \beta_b = 0$. The corresponding estimated contrast is $\widehat{\Delta}_{ab} = \hat{\beta}_a - \hat{\beta}_b$. Since both coefficients are estimated jointly within the same model, the variance of this contrast is approximated as

$$\widehat{\mathrm{Var}}(\widehat{\Delta}_{ab}) = \widehat{\mathrm{Var}}(\hat{\beta}_a) + \widehat{\mathrm{Var}}(\hat{\beta}_b) - 2\,\widehat{\mathrm{Cov}}(\hat{\beta}_a, \hat{\beta}_b). \tag{32}$$

The corresponding standard error and approximate test statistics are

$$\mathrm{se}(\widehat{\Delta}_{ab}) = \sqrt{\widehat{\mathrm{Var}}(\widehat{\Delta}_{ab})}, \qquad z_{ab} = \frac{\widehat{\Delta}_{ab}}{\mathrm{se}(\widehat{\Delta}_{ab})}, \qquad p_{ab} = 2\big[1 - \Phi(|z_{ab}|)\big]. \tag{33}$$

This tests whether the estimated difference between two jointly estimated coefficients is large relative to its uncertainty. The covariance term matters because both coefficients are obtained from the same model.

Pairwise coefficient comparisons across segments. To assess whether the same coefficient differs across two segment-specific models, I consider the null hypothesis $H_0 : \beta_h^{(s_1)} - \beta_h^{(s_2)} = 0$. The corresponding estimated contrast is $\widehat{\Delta}_{h,s_1,s_2} = \hat{\beta}_h^{(s_1)} - \hat{\beta}_h^{(s_2)}$. In the implementation, the variance of this contrast is approximated as

$$\widehat{\mathrm{Var}}(\widehat{\Delta}_{h,s_1,s_2}) \approx \widehat{\mathrm{Var}}\left(\hat{\beta}_h^{(s_1)}\right) + \widehat{\mathrm{Var}}\left(\hat{\beta}_h^{(s_2)}\right), \tag{34}$$

which corresponds to treating the two segment-specific estimates as independent. Accordingly,

$$\mathrm{se}(\widehat{\Delta}_{h,s_1,s_2}) = \sqrt{\widehat{\mathrm{Var}}(\widehat{\Delta}_{h,s_1,s_2})}, \qquad z_{h,s_1,s_2} = \frac{\widehat{\Delta}_{h,s_1,s_2}}{\mathrm{se}(\widehat{\Delta}_{h,s_1,s_2})}, \qquad p_{h,s_1,s_2} = 2\big[1 - \Phi(|z_{h,s_1,s_2}|)\big]. \tag{35}$$

These across-segment contrasts should be interpreted more cautiously than the within-segment comparisons, because they rely on the simplifying assumption of independent segment-specific estimates and therefore provide only an approximate indication of cross-segment differences.

Odds-based interpretation. In the multinomial logit model, a coefficient can also be interpreted in terms of relative odds between two alternatives. Holding all other utility components constant (*ceteris paribus*), an increase of Δx in one covariate changes those relative odds multiplicatively as

$$\frac{\mathrm{odds}^{(1)}}{\mathrm{odds}^{(0)}} = \exp(\beta_h \Delta x). \tag{36}$$

Since the destination-side features in this thesis enter the model as z-standardised variables, a one-unit increase on the model scale corresponds to a one-standard-deviation increase in the transformed covariate. Accordingly, for $\Delta x = 1$, the relative odds are multiplied by $\exp(\beta_h)$.

For variables transformed as $z(\log(1 + x))$, this one-standard-deviation change can be translated back to the original scale by reversing the transformation from a given baseline raw value x_0. Because this back-transformation is non-linear, the same one-unit increase on the model scale corresponds to different absolute changes in original units depending on the starting value. In this thesis, the raw-scale interpretation is illustrated using the median of the empirical distribution of the original variable as the reference value.

J Supplementary analysis of model results

This appendix reports additional material complementing the model results presented in Section 6.1. Its purpose is to provide a more detailed view of the empirical patterns discussed in the main text.

J.1 Distance distribution

Figure 18 compares the observed trip-distance distributions with the distributions implied by the estimated model for the four behavioural segments. The predicted distributions shown in the figure are based on the expected distances implied by the estimated utility function, whereas the observed distributions correspond to the trip distances reported in the Microcensus.

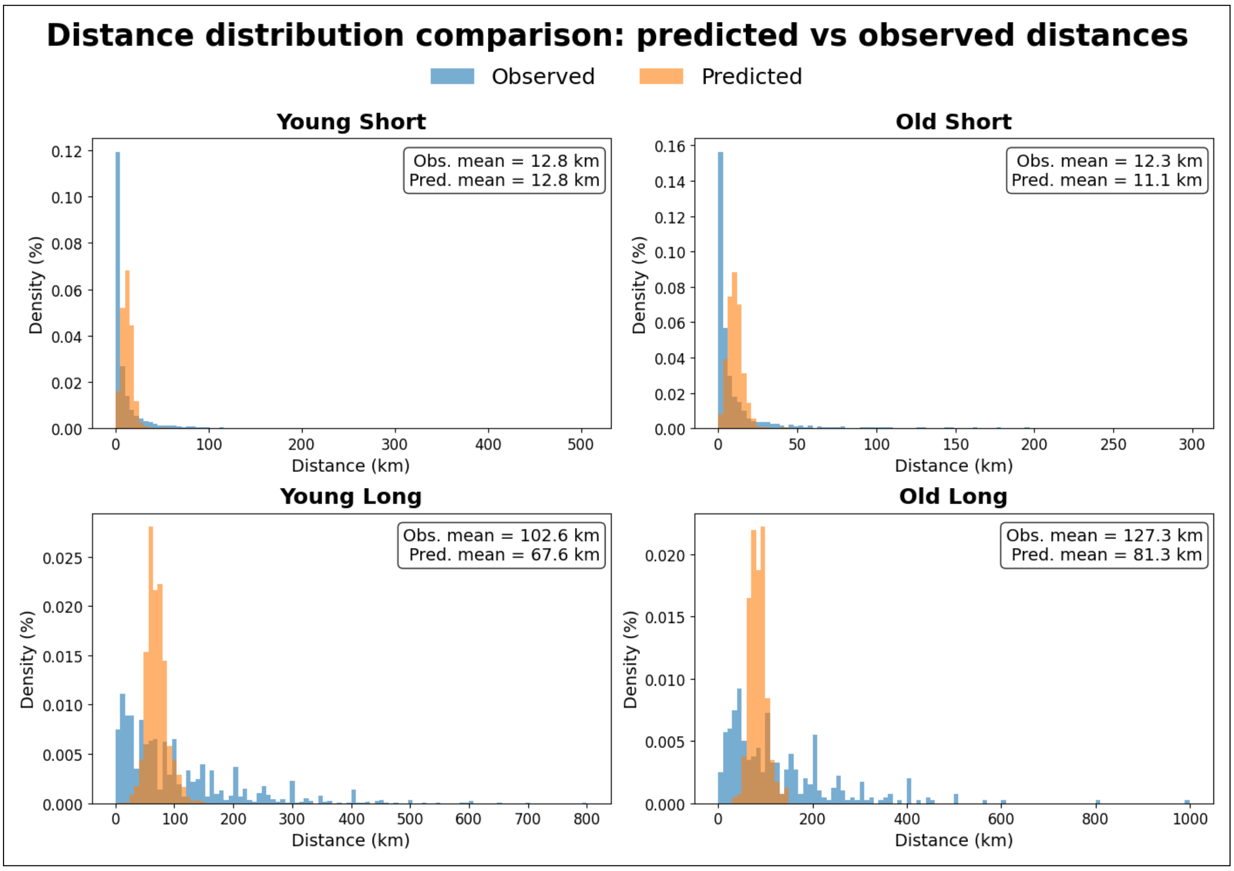


Figure 18: Comparison between observed trip-distance distributions and model-implied predicted distance distributions for the four behavioural segments.

For the two short-trip segments, the predicted distributions are broadly aligned with the observed ones. In both Young–Short and Old–Short, the model reproduces the strong concentration of trips at short distances reasonably well, and the difference between observed and predicted mean distance remains limited. Even so, the model still under-represents the right tail of the distribution, meaning that relatively longer short-leisure trips receive comparatively little probability mass. For the two long-trip segments, the comparison is more difficult to interpret. The observed distance values reported here are the original ones taken from the Microcensus. However, for long/day trips derived from *Tagesreisen*, the reported distance refers to the full trip and therefore includes both the outbound and the return leg, whereas the model-implied distance is computed as a one-way origin–destination network distance. As a result, the apparent underestimation visible in mean-distance comparisons should not be interpreted too literally. A simple heuristic is to assume approximate symmetry between outbound and return travel and therefore to divide the observed long-trip distances by two. Under such a simplification, the observed and predicted mean distances would become considerably closer, and the apparent underestimation suggested by the original values could in some cases even turn into a slight overestimation. Even with this caveat, however, the upper tail of the long-trip distance distribution remains difficult for the model to reproduce. Very long trips still appear less represented in the model-implied distributions than in the observed data. In this sense, the interpretation remains broadly similar to that of the short-trip segments: the model captures the central mass of the distribution more successfully than the extreme upper tail of realised distances.

J.2 Analysis of destination-share patterns

Figure 19 maps the estimated *attractivity component* $A_j(s)$ of destination utility for the four behavioural segments, excluding accessibility. It should therefore not be interpreted as an absolute measure of leisure attractivity, but rather as the destination-side contribution to utility within the estimated specification.

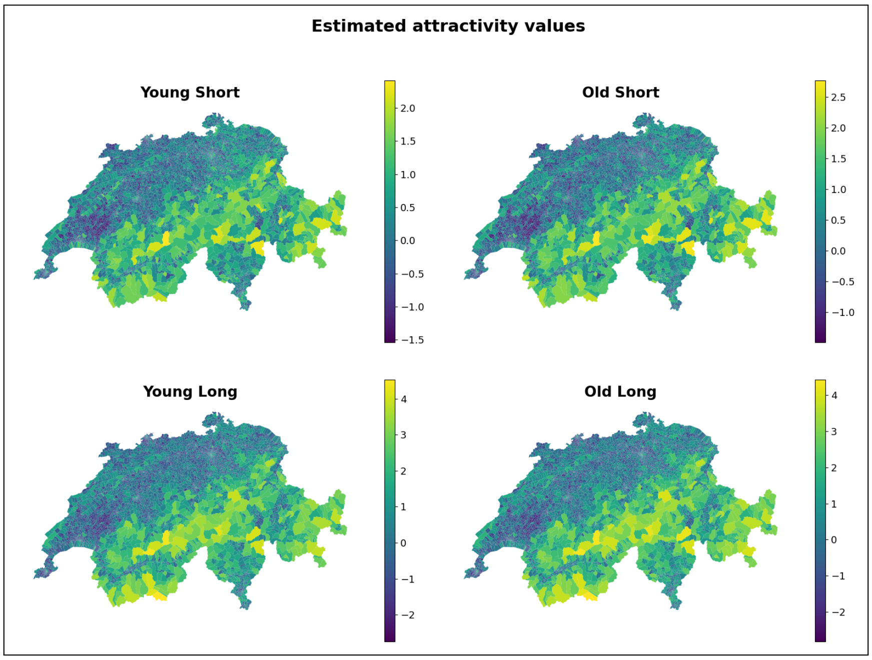


Figure 19: Estimated destination-side attractivity component $A_j(s)$ by segment.

A general pattern is that attractivity values tend to be higher along the Alpine chain and lower across large parts of the Swiss Plateau. This does not imply that plateau zones are unattractive in absolute terms. Rather, within the estimated utility function, the attractivity term tends to compensate destinations that are chosen despite relatively low accessibility, whereas highly accessible destinations can obtain high overall utility even with a more moderate destination-side contribution. In this sense, the spatial pattern of $A_j(s)$ appears broadly complementary to the accessibility pattern shown in Figure 4. This contrast is more pronounced in the two long-trip segments.

Figure 20 maps pairwise differences in the *model-implied destination shares* shown in Figure 5. Each panel isolates one segmentation dimension while keeping the other fixed. The upper row compares Young and Old within short trips (left) and within long trips (right). The lower row compares Long and Short within the Young group (left) and within the Old group (right). Positive and negative values therefore indicate which segment has the relatively higher implied share in each zone.

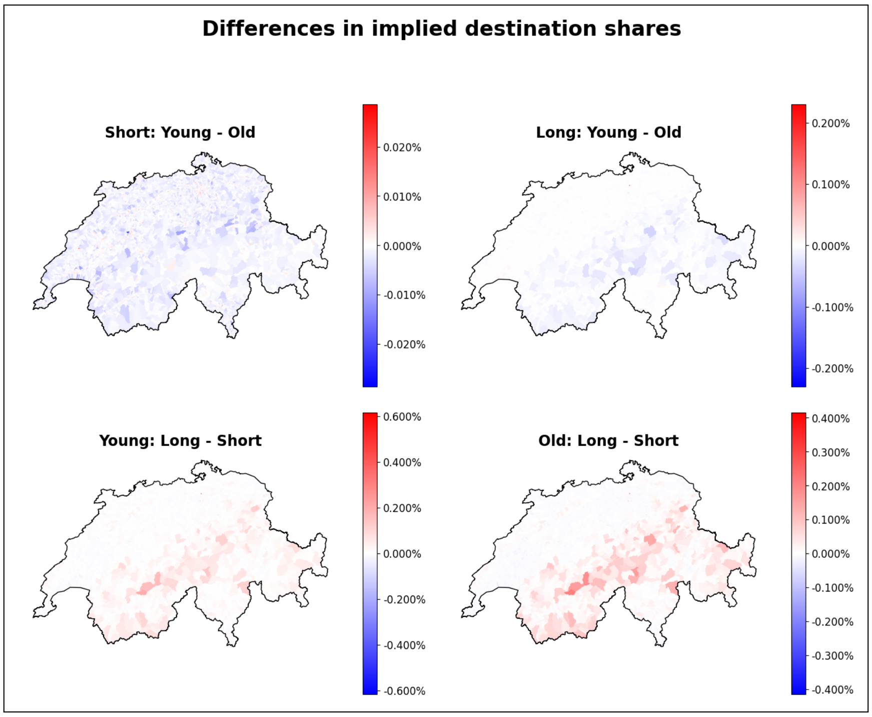


Figure 20: Pairwise differences in model-implied destination shares across segmentation dimensions.

For short trips, the Young–Old difference displays a mixed spatial pattern that is not easy to summarise visually. A tentative tendency is that older adults show relatively higher implied shares in more rural or less urban zones, whereas younger individuals tend to show relatively higher shares in larger cities such as Lausanne, Zurich, and Geneva. Overall, however, the pattern remains heterogeneous and zone-specific. Even though the absolute differences are small, the map still suggests that some age-related variation is present. For long trips, the Young–Old contrast is more structured. In particular, older adults tend to show relatively higher implied shares across several parts of the Alpine arc. The comparison between long and short trips within each age group yields a clearer pattern: Alpine and mountainous zones systematically receive relatively higher implied shares in the long-trip segments than in the short-trip segments.

A more local illustration is provided in Figure 21, which shows the model-implied destination shares for the Canton of Ticino. Ticino is a useful example because it combines a strongly Alpine northern part with a more urbanised southern part, and therefore makes it easier to visualise how destination shares vary within a relatively small but geographically diverse region.

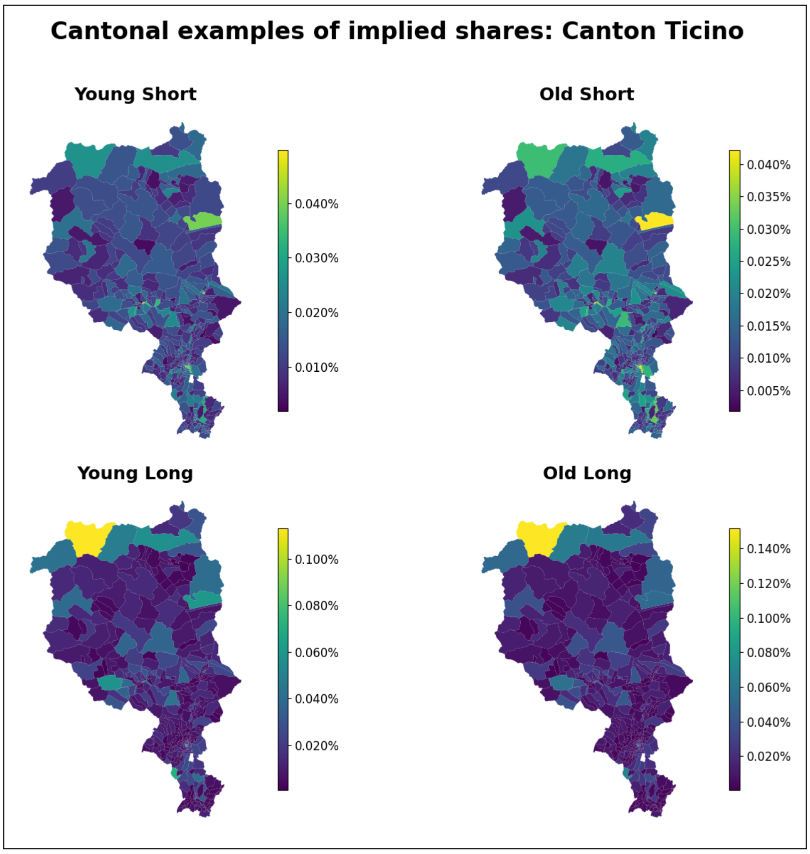


Figure 21: Model-implied destination shares in the Canton of Ticino by segment.

For short trips, the implied destination shares are relatively homogeneous across the canton. A particularly prominent hotspot appears in the Blenio Valley in northern Ticino, with another around the Gotthard area. Additional hotspots can be observed around the southern urban centres. Differences between younger and older individuals are almost imperceptible. For long trips, the spatial pattern becomes much more concentrated. The highest implied shares are located in the northern part of the canton, corresponding to the Alpine area, and especially around the Gotthard region. In this sense, the Ticino example mirrors the national-scale pattern already observed above.

J.3 Complete coefficient output and supporting correlations

This subsection reports the complete coefficient output corresponding to the compact summary shown in Table 9. For each behavioural segment, the tables report the estimated coefficient, its standard error, the associated z-statistic, the corresponding p-value, and the significance marker.

Table 22: Extended coefficient output for segment YS (Young–Short).

Variable	coef	se	z	p	stars
1. Gastronomy POIs	0.327	0.011	30.827	0.000	***
2. Resident population	-0.056	0.006	-9.633	0.000	***
3. Lake shore density	0.060	0.006	10.151	0.000	***
4. Hard outdoor POIs	-0.010	0.009	-1.126	0.260	
5. Soft outdoor POIs	-0.048	0.007	-6.479	0.000	***
6. Land-use mix	0.055	0.007	8.239	0.000	***
7. Cultural POIs	0.023	0.006	3.895	0.000	***
8. Sport POIs	0.112	0.009	12.877	0.000	***
9. Outdoor sport route length	0.200	0.014	14.382	0.000	***
10. Other leisure POIs	-0.005	0.006	-0.949	0.343	
11. Urban diversity index	0.051	0.014	3.727	0.000	***
12. Urban POIs density	0.039	0.013	3.038	0.002	**
13. Support/service amenities	0.181	0.008	21.700	0.000	***
Accessibility α_{EMU}	1.294	0.003	412.439	0.000	***

Table 23: Extended coefficient output for segment OS (Old–Short).

Variable	coef	se	z	p	stars
1. Gastronomy POIs	0.347	0.023	15.406	0.000	***
2. Resident population	-0.015	0.014	-1.055	0.292	
3. Lake shore density	0.079	0.012	6.365	0.000	***
4. Hard outdoor POIs	-0.051	0.018	-2.847	0.004	**
5. Soft outdoor POIs	-0.016	0.016	-1.021	0.307	
6. Land-use mix	0.040	0.015	2.698	0.007	**
7. Cultural POIs	0.037	0.013	2.823	0.005	**
8. Sport POIs	0.053	0.018	2.882	0.004	**
9. Outdoor sport route length	0.278	0.029	9.513	0.000	***
10. Other leisure POIs	-0.001	0.013	-0.057	0.955	
11. Urban diversity index	0.039	0.028	1.375	0.169	
12. Urban POIs density	-0.084	0.028	-3.026	0.002	**
13. Support/service amenities	0.220	0.018	12.247	0.000	***
Accessibility α_{EMU}	1.345	0.007	192.438	0.000	***

As an additional aid to interpretation, Table 26 reports the Pearson and Spearman correlations between the 13 retained destination-side variables and a destination-level accessibility summary derived from the EMU matrix. For each destination zone j, this

Table 24: Extended coefficient output for segment YL (Young–Long).

Variable	coef	se	z	p	stars
1. Gastronomy POIs	0.612	0.023	27.061	0.000	***
2. Resident population	-0.061	0.011	-5.693	0.000	***
3. Lake shore density	0.092	0.012	7.706	0.000	***
4. Hard outdoor POIs	0.074	0.015	4.821	0.000	***
5. Soft outdoor POIs	-0.194	0.016	-12.259	0.000	***
6. Land-use mix	-0.152	0.014	-10.527	0.000	***
7. Cultural POIs	-0.094	0.013	-7.385	0.000	***
8. Sport POIs	-0.060	0.018	-3.278	0.001	**
9. Outdoor sport route length	0.480	0.026	18.753	0.000	***
10. Other leisure POIs	0.090	0.011	8.262	0.000	***
11. Urban diversity index	0.178	0.031	5.791	0.000	***
12. Urban POIs density	0.037	0.027	1.380	0.168	
13. Support/service amenities	0.265	0.018	14.845	0.000	***
Accessibility α_{EMU}	0.724	0.005	135.947	0.000	***

Table 25: Extended coefficient output for segment OL (Old–Long).

Variable	coef	se	z	p	stars
1. Gastronomy POIs	0.567	0.051	11.058	0.000	***
2. Resident population	-0.037	0.026	-1.411	0.158	
3. Lake shore density	0.112	0.026	4.258	0.000	***
4. Hard outdoor POIs	0.046	0.033	1.369	0.171	
5. Soft outdoor POIs	-0.166	0.036	-4.660	0.000	***
6. Land-use mix	-0.130	0.033	-3.942	0.000	***
7. Cultural POIs	-0.102	0.029	-3.506	0.000	***
8. Sport POIs	-0.155	0.041	-3.772	0.000	***
9. Outdoor sport route length	0.533	0.057	9.363	0.000	***
10. Other leisure POIs	0.045	0.025	1.779	0.075	
11. Urban diversity index	0.289	0.070	4.132	0.000	***
12. Urban POIs density	0.036	0.060	0.600	0.548	
13. Support/service amenities	0.251	0.041	6.126	0.000	***
Accessibility α_{EMU}	0.634	0.012	51.191	0.000	***

summary is computed as the logsumexp across all origins,

$$\mathrm{EMU}_j^{\mathrm{dest}} = \log\left(\sum_i \exp(\mathrm{EMU}_{ij})\right) \tag{37}$$

This provides a synthetic indication of how accessible a destination is from the full set of origins. The reported correlations are not used for estimation, but help assess whether a destination attribute tends to align with accessibility, oppose it, or capture a relatively distinct dimension. To complement this, Table 27 reports the pairwise Pearson correlations among the 13 retained destination-side variables. The variable indices used in Table 27 correspond to the numbering reported in Table 26.

Table 26: Correlation between the retained destination-side variables and destination-side accessibility derived from the EMU matrix.

Index	Variable	Pearson	Spearman
1	Gastronomy POIs	0.022	0.044
2	Resident population	0.163	0.237
3	Lake shore density	-0.203	-0.238
4	Hard outdoor POIs	-0.491	-0.411
5	Soft outdoor POIs	-0.066	-0.032
6	Land-use mix	-0.286	-0.261
7	Cultural POIs	0.073	0.079
8	Sport POIs	0.002	-0.002
9	Outdoor sport route length	-0.762	-0.709
10	Other leisure POIs	0.121	0.114
11	Urban diversity index	0.035	0.062
12	Urban POIs density	0.655	0.680
13	Support/service amenities	-0.226	-0.171

Table 27: Pairwise Pearson correlations among the 13 retained destination-side variables. Variable indices correspond to Table 26.

Idx	1	2	3	4	5	6	7	8	9	10	11	12	13
1	**1.00**												
2	0.11	**1.00**											
3	0.17	0.00	**1.00**										
4	0.23	0.00	0.23	**1.00**									
5	0.22	0.35	0.21	0.33	**1.00**								
6	-0.04	0.17	0.26	0.16	0.24	**1.00**							
7	0.44	0.06	0.08	0.09	0.15	-0.03	**1.00**						
8	0.17	0.35	0.12	0.15	0.45	0.31	0.14	**1.00**					
9	0.12	-0.02	0.23	0.67	0.27	0.33	-0.01	0.17	**1.00**				
10	0.10	0.02	-0.01	-0.01	0.03	-0.06	0.07	0.01	-0.08	**1.00**			
11	0.70	0.23	0.17	0.17	0.34	0.11	0.53	0.55	0.12	0.28	**1.00**		
12	0.43	0.09	-0.14	-0.35	-0.03	-0.33	0.33	0.15	-0.65	0.22	0.44	**1.00**	
13	0.51	0.12	0.24	0.43	0.42	0.21	0.34	0.32	0.41	0.04	0.49	0.00	**1.00**

K Excursus on finer age segmentation

As an additional sensitivity exercise, I re-estimate the short-trip specification using a finer age segmentation, splitting the broad short-trip population into six age bins. The aim is to assess whether the main short-trip results conceal additional age-related heterogeneity, and whether a more granular segmentation improves predictive performance. Table 28 reports the full-choice-set out-of-sample metrics for each age-bin-specific model, together with the corresponding pooled short-trip model estimated without age segmentation. Figure 22 complements the metric comparison by showing how the estimated destination-side coefficients evolve across the finer age classes.

Table 28: Out-of-sample full-choice-set performance under finer age segmentation for short leisure trips.

Group	N	WP	**NLL**	**MRR**	R^2	**Pearson**	**Spearman**	**JS**
6–15	1866	1577.85	4.136	0.284	0.540	0.492	0.366	0.347
16–29	2206	2613.57	5.144	0.174	0.427	0.509	0.348	0.343
30–49	3146	3623.98	5.019	0.208	0.441	0.466	0.367	0.304
50–64	2776	2518.28	4.810	0.218	0.465	0.442	0.353	0.321
65–79	2336	1857.90	4.772	0.233	0.469	0.480	0.355	0.332
80+	434	504.84	4.494	0.254	0.500	0.374	0.224	0.499
Short only	12764	12696.417	4.854	0.217	0.455	0.644	0.511	0.138

The model performs best for the youngest group and remains comparatively strong for the older age classes, especially 50–64, 65–79, and 80+. By contrast, predictive performance is weaker for the intermediate adult groups, particularly 16–29 and 30–49. A plausible interpretation is that these middle-age groups are behaviourally more heterogeneous, because their leisure destination choices may depend more strongly on additional dimensions that are not explicitly modelled here. As already suggested by the literature discussed in Section 3.3, mobility behaviour may vary not only by age, but also by socio-economic status, employment situation, and life-course context. Nevertheless, the finer age-bin exercise further confirms the presence of heterogeneity in leisure destination choice. At the same time, it suggests that a more detailed age split alone may still be insufficient to capture the full complexity of the underlying behaviour. However, the comparatively good performance of the 65+ groups under a finer age partition still supports the relevance of the age-based segmentation adopted in the main analysis.

Some destination-side coefficients remain close to zero and vary only little across the age classes. This is the case, for example, for *Cultural POIs (7)* and *Land-use mix (6)*, whose estimated effects appear fairly stable throughout the finer segmentation. Other variables display more visible heterogeneity. This applies in particular to *Gastronomy POIs (1)*, *Urban POIs density (12)*, and *Outdoor sport route length (9)*, whose coefficients vary more substantially across age groups. For some variables, differences emerge

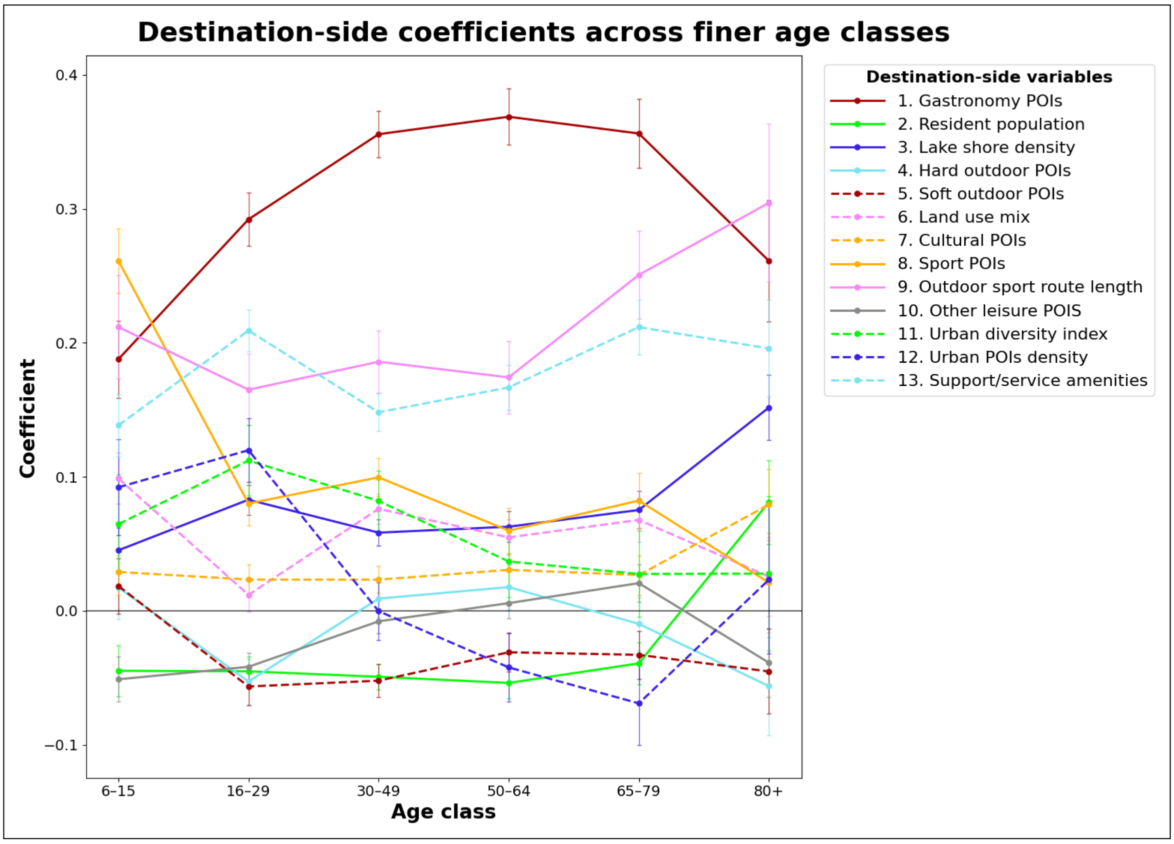


Figure 22: Estimated destination-side coefficients across the finer short-trip age classes. The figure reports the coefficients together with approximate ±1 standard-error bands.

only for specific classes: *Sport POIs (8)*, for instance, stand out mainly for the youngest group. A further noteworthy pattern concerns *resident population*, whose coefficient becomes positive for the 80+ group and already shows a slight upward shift for the 65–79 class. More generally, the main differences across age groups appear to be driven less by any single variable in isolation than by the overall combined structure of destination attractivity, consistent with the broader patterns already observed in the main analysis.

These results are only exploratory and should not be over-interpreted, especially because some age bins contain substantially fewer observations than others. Even so, the exercise is useful in two ways. First, it reinforces the idea that segmentation is valuable in leisure destination-choice modelling, since relevant behavioural differences do emerge once more homogeneous subgroups are considered. Second, it suggests that further improvements may require going beyond age alone and incorporating additional segmentation dimensions or explanatory variables. In this sense, the finer age-bin analysis provides a useful indication for future work: with richer data and more detailed segmentation variables, even stronger behavioural insights could likely be obtained.

L Robustness analysis metrics

This appendix reports the detailed quantitative results underlying the robustness analysis discussed in Section 6.3. Table 29 reports, for each alternative specification, the difference in out-of-sample performance relative to the corresponding baseline model within the same segment. The table is organised hierarchically by robustness family $R1$–$R7$, following the structure introduced in Section 5.5. In addition, Table 30 reports the differences in estimated coefficients between the specification excluding *Visits* observations and the corresponding baseline model. For interpretative convenience, negative values indicate an improvement for ΔNLL and ΔJS, whereas positive values indicate an improvement for ΔMRR, ΔR^2, and $\Delta\rho_S$. Values close to zero therefore indicate that the corresponding robustness check leaves the baseline conclusions largely unchanged.

Table 29: Out-of-sample performance differences relative to the corresponding baseline specification within the same segment (Δ = modified specification − baseline).

Specification	**Segment**	ΔNLL	ΔMRR	ΔR^2_{McF}	$\Delta\rho_S$	ΔJS
R1: Preprocessing choices						
Excluding Liechtenstein	YS	0.0481	-0.0056	-0.0055	0.0047	0.0029
	OS	-0.1361	0.0114	0.0151	-0.0138	0.0102
	YL	-0.0367	0.0056	0.0039	-0.0105	0.0031
	OL	0.0961	-0.0017	-0.0109	0.0004	-0.0017
Raw instead of $\log(1 + x)$	YS	0.0136	-0.0047	-0.0015	-0.0079	0.0035
	OS	0.0245	-0.0020	-0.0027	-0.0080	0.0048
	YL	0.1177	-0.0013	-0.0131	-0.0770	0.0136
	OL	0.1741	-0.0058	-0.0194	-0.0463	0.0119
No standardisation	YS	-0.0011	-0.0021	0.0001	0.0012	0.0000
	OS	0.0000	0.0008	0.0000	-0.0004	0.0006
	YL	-0.0002	-0.0012	0.0000	0.0001	0.0000
	OL	-0.0085	0.0001	0.0009	-0.0014	-0.0031
Density-based amenities	YS	0.0104	-0.0024	-0.0012	-0.0105	0.0033
	OS	0.0067	0.0000	-0.0007	-0.0026	0.0025
	YL	0.0373	-0.0004	-0.0041	-0.0133	0.0071
	OL	0.0256	-0.0056	-0.0028	-0.0111	0.0006
Four-variable specification	YS	0.0056	-0.0001	-0.0006	-0.0017	0.0011
1., 8., 9., 13.	OS	0.0065	-0.0123	-0.0007	-0.0033	0.0012
	YL	0.0575	-0.0003	-0.0064	0.0126	-0.0009
	OL	0.0211	-0.0015	-0.0023	0.0090	-0.0032
No visits	YS	-0.1850	0.0125	0.0206	0.0115	0.0240

Continued on next page

Table 29 continued from previous page

Specification	Segment	ΔNLL	ΔMRR	$\Delta R^2_{\mathrm{McF}}$	$\Delta\rho_S$	ΔJS
	OS	-0.1935	0.0099	0.0215	-0.0008	0.0259
	YL	-0.2423	0.0203	0.0270	0.0015	0.0108
	OL	-0.0743	-0.0064	0.0083	-0.0164	0.0240
Drop 1.	YS	0.0141	0.0011	-0.0016	-0.0128	0.0034
	OS	0.0138	-0.0029	-0.0015	-0.0059	0.0034
	YL	0.0605	0.0007	-0.0067	-0.0249	0.0108
	OL	0.0621	-0.0059	-0.0069	-0.0145	0.0052
Drop 2.	YS	0.0014	-0.0006	-0.0002	-0.0002	0.0001
	OS	0.0003	-0.0043	0.0000	0.0000	0.0005
	YL	0.0038	-0.0008	-0.0004	-0.0030	0.0016
	OL	0.0012	-0.0008	-0.0001	-0.0033	-0.0004
Drop 3.	YS	0.0014	-0.0003	-0.0002	-0.0002	0.0007
	OS	0.0035	-0.0035	-0.0004	-0.0012	0.0010
	YL	0.0013	0.0023	-0.0001	-0.0017	0.0014
	OL	0.0111	-0.0054	-0.0012	0.0142	-0.0025
Drop 4.	YS	0.0006	-0.0027	-0.0001	-0.0005	0.0004
	OS	0.0006	-0.0010	-0.0001	-0.0012	0.0005
	YL	0.0017	0.0000	-0.0002	-0.0045	0.0014
	OL	0.0057	-0.0046	-0.0006	-0.0086	0.0016
Drop 5.	YS	0.0001	-0.0010	0.0000	0.0004	-0.0002
	OS	0.0010	-0.0093	-0.0001	-0.0010	0.0005
	YL	0.0042	0.0015	-0.0005	-0.0089	0.0015
	OL	0.0079	-0.0055	-0.0009	0.0027	-0.0003
Drop 6.	YS	0.0012	-0.0001	-0.0001	-0.0005	0.0001
	OS	0.0015	-0.0043	-0.0002	-0.0017	0.0000
	YL	0.0073	0.0019	-0.0008	-0.0004	0.0009
	OL	0.0080	-0.0003	-0.0009	0.0163	-0.0031
Drop 7.	YS	0.0001	-0.0008	0.0000	-0.0001	0.0002
	OS	0.0017	-0.0020	-0.0002	-0.0025	0.0001
	YL	0.0072	0.0012	-0.0008	-0.0030	0.0024
	OL	0.0251	-0.0017	-0.0028	0.0005	-0.0003
Drop 8.	YS	0.0033	-0.0038	-0.0004	-0.0008	0.0007
	OS	0.0018	-0.0023	-0.0002	-0.0006	0.0002
	YL	0.0003	-0.0014	0.0000	-0.0008	0.0017
	OL	-0.0024	-0.0034	0.0003	0.0203	-0.0027

Continued on next page

Table 29 continued from previous page

Specification	Segment	ΔNLL	ΔMRR	$\Delta R^2_{\mathrm{McF}}$	$\Delta\rho_S$	ΔJS
Drop 9.	YS	0.0002	0.0000	0.0000	0.0000	-0.0004
	OS	0.0084	-0.0098	-0.0009	-0.0074	0.0014
	YL	0.0255	0.0015	-0.0028	-0.0129	0.0022
	OL	0.0502	-0.0046	-0.0056	0.0195	-0.0014
Drop 10.	YS	0.0003	-0.0038	0.0000	0.0006	-0.0002
	OS	0.0000	-0.0087	0.0000	-0.0002	0.0001
	YL	0.0049	0.0012	-0.0005	-0.0060	0.0022
	OL	0.0143	-0.0054	-0.0016	0.0119	0.0024
Drop 11.	YS	0.0005	0.0003	-0.0001	0.0004	0.0001
	OS	-0.0004	-0.0083	0.0000	0.0002	-0.0001
	YL	0.0007	-0.0010	-0.0001	-0.0043	0.0021
	OL	-0.0040	-0.0011	0.0005	0.0161	-0.0021
Drop 12.	YS	-0.0001	-0.0016	0.0000	0.0006	-0.0001
	OS	-0.0008	-0.0061	0.0001	0.0013	-0.0002
	YL	-0.0006	0.0003	0.0001	-0.0017	0.0013
	OL	0.0068	-0.0039	-0.0008	0.0156	-0.0001
Drop 13.	YS	0.0061	0.0017	-0.0007	-0.0103	0.0014
	OS	0.0037	-0.0022	-0.0004	-0.0056	0.0017
	YL	0.0049	0.0023	-0.0005	0.0056	0.0044
	OL	0.0362	-0.0082	-0.0040	0.0162	0.0019
R2: Accessibility component						
Log-beeline distance	YS	0.0727	0.0410	-0.0081	0.0178	-0.0044
	OS	0.0792	0.0553	-0.0088	0.0094	-0.0078
	YL	0.0951	0.0009	-0.0106	-0.0059	0.0041
	OL	0.1243	0.0024	-0.0138	0.0029	0.0006
Log travel-distance skim	YS	0.0030	0.0020	-0.0003	0.0044	-0.0005
	OS	-0.0112	-0.0041	0.0012	0.0028	-0.0010
	YL	0.0776	-0.0037	-0.0086	-0.0018	0.0023
	OL	0.0892	-0.0054	-0.0099	0.0023	0.0001
R3: Sampled choice-set construction						
K = 100	YS	0.0014	-0.0030	-0.0002	-0.0022	-0.0003
	OS	0.0051	-0.0018	-0.0006	-0.0051	0.0018
	YL	0.0045	0.0005	-0.0005	0.0018	-0.0016
	OL	-0.0014	-0.0022	0.0002	0.0009	0.0008
K = 300	YS	0.0005	0.0014	-0.0001	-0.0006	-0.0005

Continued on next page

Table 29 continued from previous page

Specification	Segment	ΔNLL	ΔMRR	ΔR^2_{McF}	$\Delta\rho_S$	ΔJS
	OS	0.0016	0.0044	-0.0002	0.0018	0.0006
	YL	0.0007	0.0007	-0.0001	0.0007	-0.0003
	OL	0.0014	-0.0020	-0.0002	0.0002	0.0003
K = 500	YS	-0.0002	0.0025	0.0000	0.0000	-0.0005
	OS	0.0030	-0.0001	-0.0003	-0.0001	-0.0006
	YL	0.0004	0.0019	0.0000	0.0027	-0.0017
	OL	0.0000	-0.0013	0.0000	-0.0002	0.0005
K = 800	YS	-0.0001	0.0000	0.0000	0.0001	-0.0006
	OS	-0.0002	-0.0044	0.0000	-0.0004	-0.0004
	YL	0.0003	0.0008	0.0000	-0.0001	0.0000
	OL	-0.0004	-0.0036	0.0000	-0.0002	0.0005
K = 2000	YS	0.0015	0.0030	-0.0002	-0.0032	0.0014
	OS	0.0026	-0.0010	-0.0003	-0.0016	0.0006
	YL	0.0000	0.0008	0.0000	-0.0003	-0.0001
	OL	-0.0014	-0.0035	0.0002	-0.0001	0.0003
K = 3000	YS	0.0008	0.0034	-0.0001	0.0000	0.0006
	OS	-0.0010	0.0013	0.0001	0.0018	-0.0008
	YL	0.0005	0.0019	-0.0001	0.0009	-0.0010
	OL	-0.0009	-0.0003	0.0001	0.0000	0.0004
EMU-weighted sampling	YS	0.0293	0.0023	-0.0033	0.0078	-0.0021
	OS	0.0291	-0.0050	-0.0032	0.0058	-0.0064
	YL	0.0366	0.0007	-0.0041	0.0011	-0.0010
	OL	0.0186	-0.0028	-0.0021	0.0001	-0.0037
R4: Out-of-sample split design						
By-origin split	YS	-0.0074	-0.0030	0.0008	0.0167	0.0020
	OS	-0.1042	0.0126	0.0116	0.0108	-0.0005
	YL	0.0195	0.0045	-0.0022	0.0051	-0.0072
	OL	0.2147	-0.0099	-0.0239	-0.0106	-0.0021
R5: Survey-year sensitivity						
2021 only	YS	0.0615	0.0000	-0.0068	-0.0742	0.0866
	OS	0.0852	-0.0070	-0.0095	-0.0623	0.0767
	YL	-0.0458	-0.0006	0.0051	-0.0219	0.0384
	OL	-0.0808	-0.0031	0.0090	-0.0110	0.0233
2015 only	YS	-0.0647	-0.0009	0.0072	-0.0626	0.0743
	OS	-0.0848	0.0087	0.0094	-0.0521	0.0800

Continued on next page

Table 29 continued from previous page

Specification	**Segment**	ΔNLL	ΔMRR	$\Delta R^2_{\mathrm{McF}}$	$\Delta\rho_S$	ΔJS
	YL	0.0374	0.0036	-0.0042	-0.0430	0.0470
	OL	0.0895	-0.0017	-0.0100	-0.0485	0.0492
R6: Regularisation						
L1 regularisation	YS	-0.0004	-0.0024	0.0000	0.0006	0.0000
($\lambda \sum_p \lvert\beta_p\rvert$)	OS	-0.0009	0.0066	0.0001	0.0003	-0.0001
	YL	0.0001	0.0015	0.0000	0.0000	0.0000
	OL	0.0000	-0.0017	0.0000	0.0000	0.0000
L2 regularisation	YS	-0.0002	-0.0002	0.0000	0.0001	0.0003
($\lambda \sum_p \beta_p^2$)	OS	-0.0003	0.0030	0.0000	0.0001	0.0003
	YL	0.0017	0.0017	-0.0002	0.0018	-0.0012
	OL	0.0000	0.0008	0.0000	0.0000	0.0000
R7: Spatial spillovers						
With spillovers	YS	-0.0068	0.0027	0.0008	0.0031	-0.0014
	OS	-0.0034	-0.0060	0.0004	0.0011	-0.0011
	YL	-0.0997	0.0015	0.0111	0.0068	-0.0052
	OL	-0.0791	-0.0017	0.0008	0.0074	-0.0077

Table 30: Coefficient differences between the *No visits* specification and the baseline full sample (Δ = No visits − Full). Positive values indicate that the coefficient becomes larger when socially motivated visits are excluded, whereas negative values indicate that it becomes smaller.

Variable	**YS**	**OS**	**YL**	**OL**
1. Gastronomy POIs	0.0879	0.0497	0.0233	0.0216
2. Resident population	-0.0802	-0.0464	-0.0364	-0.0282
3. Lake shore density	0.0172	0.0475	0.0349	0.0363
4. Hard outdoor POIs	0.0143	0.0372	0.0655	0.0828
5. Soft outdoor POIs	-0.0128	-0.0446	0.0118	-0.0044
6. Land-use mix	0.0320	-0.0002	0.0235	0.0489
7. Cultural POIs	0.0073	-0.0036	-0.0277	0.0014
8. Sport POIs	0.0328	0.0129	-0.0434	-0.0334
9. Outdoor sport route length	-0.0250	0.0863	0.1182	0.0006
10. Other leisure POIs	-0.0124	-0.0260	0.0125	-0.0198
11. Urban diversity index	0.0474	0.0163	0.0228	0.0591
12. Urban POIs density	0.0090	0.1146	0.1161	-0.0625
13. Support/service amenities	0.0214	-0.0151	-0.0417	0.0778
Accessibility α_{EMU}	0.0404	0.0607	0.0225	0.0497

M Earlier coefficient set used in SIMBA integration

This appendix reports the earlier coefficient set that was used for the SIMBA MOBi integration for the two short-trip segments (Table 31). To assess whether these differences are likely to matter in practice for the SIMBA implementation, Tables 32 and 33 additionally report the correlations between the earlier and corrected versions of, respectively, the destination attractivity indices and the resulting utility matrices.

Table 31: Earlier estimated coefficients used for the SIMBA MOBi integration, for the short-trip segments.

Variable	**YS**	**OS**
1. Gastronomy POIs	0.205***	0.122***
2. Resident population	0.161***	0.264***
3. Lake shore density	0.022***	0.050***
4. Hard outdoor POIs	-0.017*	-0.012
5. Soft outdoor POIs	-0.006	0.014
6. Land-use mix	0.017***	0.017
7. Cultural POIs	0.017***	0.003
8. Sport POIs	0.097***	0.050***
9. Outdoor sport route length	0.147***	0.206***
10. Other leisure POIs	0.002	0.013
11. Urban diversity index	0.009	-0.019
12. Urban POIs density	-0.058***	-0.070***
13. Support/service amenities	0.085***	0.112***
Accessibility α_{EMU}	1.328***	1.408***

Table 32: Correlation between earlier and current attractivity indices.

Segment	**Pearson**	**Spearman**
YS	0.882	0.895
OS	0.803	0.828

Table 33: Correlation between earlier and current utility matrices.

Segment	**Pearson**	**Spearman**
YS	0.989	0.984
OS	0.985	0.977

N SIMBA integration: spatial distribution of destination shares

This appendix reports the spatial distribution of destination shares extracted from the SIMBA MOBi simulation outputs for the two main setups considered in the integration analysis: the *Reference* specification (Figure 23) and the *Proposed* specification (Figure 24). Shares are shown separately for the four operational segments (Young–Short, Old–Short, Young–Long, Old–Long). The figures are intended as a descriptive complement to the correlation-based benchmarking reported in Section 6.4.

A small number of extreme outliers emerged in the *Proposed* simulation outputs. For cartographic visualisation only, these values were capped at 0.3 in order to avoid compressing the remaining distribution and to improve the readability of the maps.

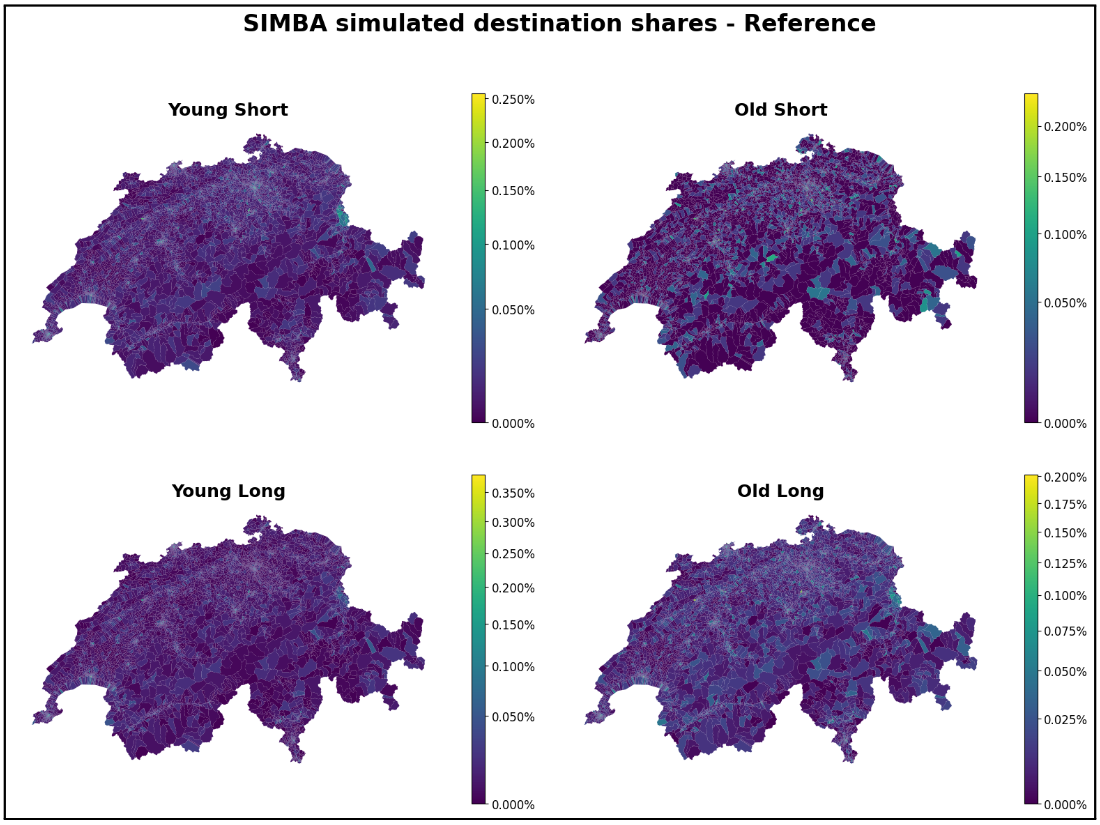


Figure 23: Spatial distribution of simulated leisure destination shares in the SIMBA MOBi *Reference* setup, shown separately for the four operational segments.

The maps should be interpreted with caution and mainly as visual support for the quantitative comparisons discussed in the main text. Even after capping the extreme values, the *Proposed* specification appears less spatially smooth and more concentrated in a limited number of zones than the *Reference* specification. At the same time, the long-trip segments do not show a spatial pattern as clearly concentrated as in the offline implied-share results. For the short-trip segments, however, the spatial patterns generated by the *Proposed* specification remain broadly plausible. Overall, the figures are best read as exploratory evidence complementing the correlation-based analysis rather than as a

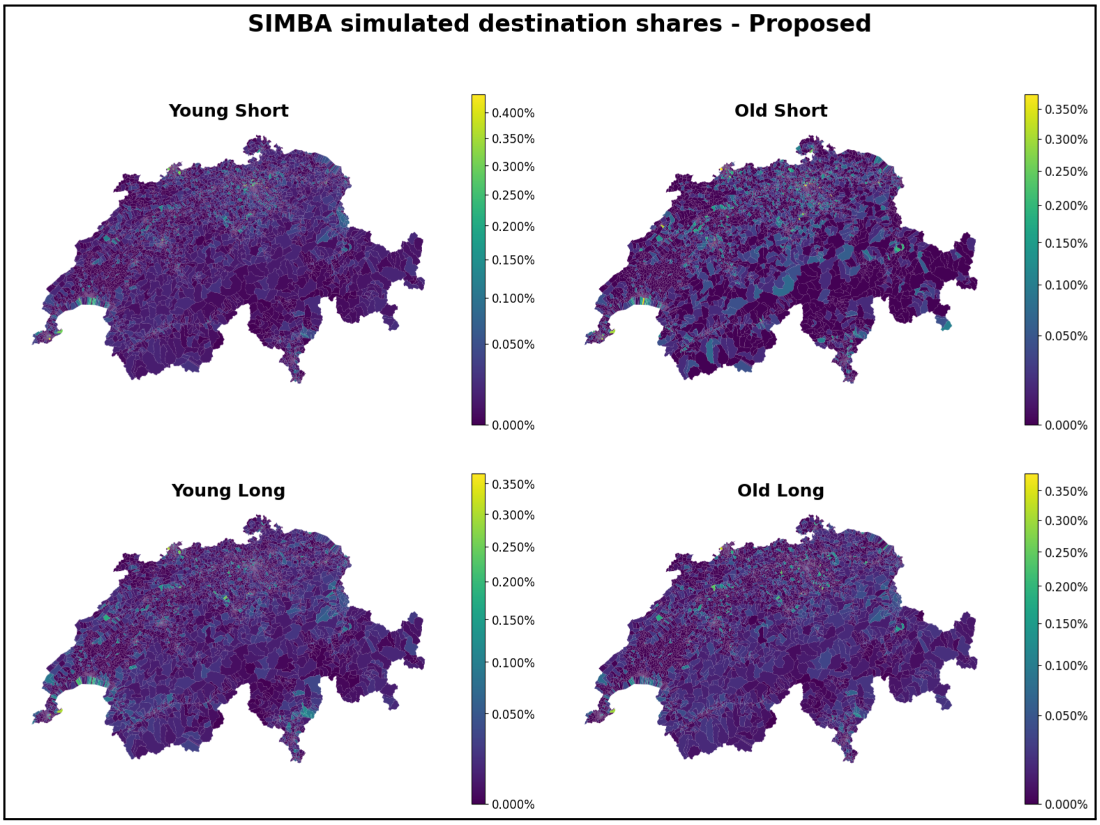


Figure 24: Spatial distribution of simulated leisure destination shares in the SIMBA MOBi *Proposed* setup, shown separately for the four operational segments.

basis for strong stand-alone conclusions.

Personal Declaration

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

In addition, I declare that ChatGPT (OpenAI) was used as a supporting tool during the preparation of this thesis. It was used for coding support, clarification of difficult concepts, consultation and discussion, translation from Italian into English, and revision of grammar and linguistic expression. The use of ChatGPT was limited to supportive and editorial functions. All substantial intellectual contributions, including the development of the research, the choice and assessment of the methods, the evaluation and interpretation of the results, the production of the final outputs, and the underlying ideas of the thesis, remain entirely my own. I assume full responsibility for all statements, analyses, formulations, and conclusions contained in this document.

Zurich, April 24, 2026

S. Benzoni

Stefano Benzoni