Individual habits: a novel operational metric and its potential as a lever of overcoming modal shift resistance in Lausanne, Switzerland

Zachary Hansen, Vincent Kaufmann[*], Jules Grandvillemin[*]

Abstract

The purpose of this paper is to investigate the sources and consequences of mode choice habits in Lausanne, Switzerland, and assess how they can help explain the gaps between current planning tools and actual public transport usage. By analyzing combined panel questionnaire and GPS tracking data collected over the course of four weeks from the Panel Lemanique, the paper establishes a human-centric definition of mode choice habits from literature and proposes a method for quantifying and characterizing the strength of a mode choice habit anchored in this definition using a Hidden Markov Model. This data is subsequently corresponded with an existing spatial index of network service quality to identify spatial relationships between network service quality and mode choice habits, before finally being used to calibrate a logit mode choice model for Lausanne. The results of these analyses are used as a first step toward understanding the role of habits in closing the gap between planning tools and actual behavior, and to recommend improvements to current planning tools to better account for mode choice habits in future planning decisions.

Keywords: Mode choice habits; Hidden Markov Model; modal shift; GPS tracking; mode choice models

Acknowledgements: The data and analysis in this paper were supported by the Ville de Lausanne Service de la Mobilité et Planification des Infrastructures (MAP), les Transports Publics Lausannois (t-l), la Direction générale de la mobilité et des routes (DGMR) du Canton de Vaud, and the Laboratory of Urban Sociology (LaSUR) at EPFL.

[*]Laboratory of Urban Sociology (LaSUR), Ecole Polytechnique Fédérale de Lausanne (EPFL), Station 16, Bâtiment BP, CH-1015 Lausanne, Switzerland

Contents

List of Figures

List of Tables

1 Introduction

1.1 Background

Transportation is the second highest-emitting economic sector in Europe, with over 8 billion tons of carbon in 2019, of which nearly half comes from cars [1]. This means that in the context of sustainable development, mobility is one of the strongest levers of action. This has led to a major planning push in recent decades to kick-start a modal shift away from car-dependence in order to meet emissions targets.

However, despite the modal shift lying at the forefront of many planning agencies' agendas, in many parts of the world, enacted measures intended to drive the modal shift appear to be having little impact on car use. Clearly, policymakers want this modal shift to happen, and they are (hopefully) not enacting policies without some sort of justification, which suggests that either the tools used to inform this decision-making are imperfect, or that the measures devised using them are too broad to effectively persuade the most impactful population segments. In fact, we do see that many transportation planning tools arising from economic disciplines (service metrics, models of accessibility, numerical mode choice models) are based fundamentally on an assumption of perfect rationality, but we know that humans do not always behave rationally: recent research has shown that these tools are often not in line with how individuals actually behave. Yet as policies are often designed using models based on this assumption of a perfect, omniscient rationality, we cannot be surprised when humans do not behave according to a behavioral assumption that is demonstrably false.

An explanation for this apparent irrationality which has gained traction in the past decades is that of modal habits: a propensity for a specific behavior to be activated in a certain context. In the context of daily mobility, we can become habituated to using a particular transportation mode to fulfill our daily mobility requirements, and as the initial mode choice is reinforced, the propensity for using this mode only increases. And in daily life, it is rare to do something for the first time: repetition is the rule, rather than the exception. The notion of "choice" is thus almost an illusion; we do not wake up every day and "choose" to take the car to work, but rather do so because it's a part of our (individual or societal) routine.

In practice, however, despite habits playing a major role in our daily mobility, planners tend to give them little attention, focusing instead on tools derived from the assumption of *homo oeconomicus*. But with our mobility patterns changing dramatically as our cities' sprawl accelerates and our homes become increasingly separated from work, the highly irrational pattern of car dependence is more firmly anchored in our society than ever before. Thus by understanding the causes and effects of modal habits, we are better able to incorporate them into mobility planning, which in turn offers a more direct lever of action for driving the modal shift–even allowing us to tailor strategies to specific segments of the population who are most strongly habituated to the automobile.

Here we provide an operational definition of habits, underscoring them as an intermediate actor between reflex and rationality, with a strong influence on daily actions such as mode "choice", that we can use as a tool to improve planning decisions. This work will serve as a first step to build on existing, imperfect planning tools by bringing focus to the role of individual mode choice habits in contributing to resistance to the modal shift toward public transport in Lausanne. As an outcome, three categories of existing planning tools will be evaluated, and new planning tools (spatial indices and mode choice models) will be proposed to help public transport operators understand and quantify the spatial distribution of these individual habits and resistances on a population-wide scale, as well as make more targeted network improvements to kick-start the modal shift for both highly car-dependent users and other users resistant to using public transport.

The Lausanne-agglomeration specifically was selected for a number of reasons, which will be presented in the rest of this section.

1.2 Context

Lausanne is a car-dependent city. According to the latest estimates from the City, rates of car ownership have remained approximately stable over the past few decades relative to the other major Swiss agglomerations, declining only 7.5% since 2006 relative to an average of almost 20% over the same period in Geneva, Bern, Zurich, and Basel [2]. Journeys by car in Lausanne remain frequent—70% of all journeys [3]—and short—more than 50% are for journeys less than 5 km [2]. Further evidence of automobile dependence is the observation that despite only comprising 50% of all drivers in Lausanne, daily car users generate 80% of the automobile traffic [4]. These numbers are significantly higher than the other two large agglomerations studied (Geneva and Bern, each with around 30% of drivers who drive daily generating 70% of the traffic) and comparable to the medium-sized urban centers in more rural parts of the Canton of Vaud. Finally, despite a public transit network with multiple bus and metro lines at very high frequency with a very high spatial coverage, convenient connectivity, and a 30% increase in regional public transport capacity from 2010-2018 [5], public transport pass ownership in Lausanne is among the lowest of all Swiss agglomerations [6].

This is reminiscent of the paradox proposed by Thomas Buhler in an examination of the Lyon public transport network; despite public transport being more affordable, and in many cases even more accessible than automobile use, automobile use persists [7, p.43-46]. Buhler's conclusion: the classical—orthodox, even—assumption in traditional planning tools is fundamentally misaligned with reality.

An in-depth diagnostic study of mode choice rationales in Lausanne and key urban areas in French-speaking Switzerland performed at EPFL's Laboratory of Urban Sociology (LaSUR) revealed that mode choice in these urban areas had strong socio-spatial characteristics [5]. The survey of around 2,000 active individuals living in close proximity to public transit found connections between mode predispositions and sociodemographic variables of age, sex, and level of education. In addition, automobile dependence was found to generally increase with distance from the city center, and predispositions for alternatives exhibiting the inverse trend.

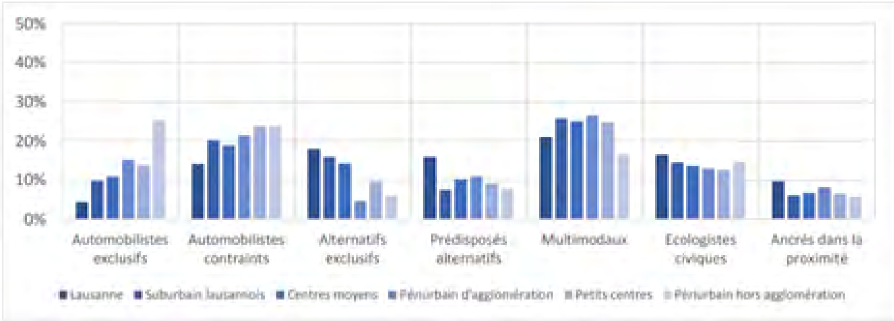


Figure 1-1: Distribution of modal predispositions in Lausanne, by distance from the city center, from [5].

At first glance, these patterns appear to be simply a consequence of city size. After all, Zurich is a larger city with less automobile use, and Lugano is a smaller city with more automobile use. However, plotting the car use values against both population and population density (Figure 1-2), along with a power-series regression, we can see quite clearly that Lausanne is an outlier, and that city size alone is not enough to explain Lausanne as an outlier. In fact, examining influence coefficients for each of the cities on the regression coefficient, we observe that Lausanne is the second-most negatively influential data point on the regression fit (behind Winterthur).

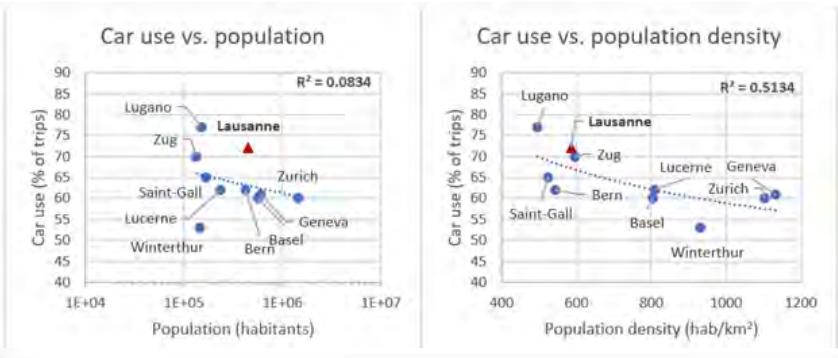
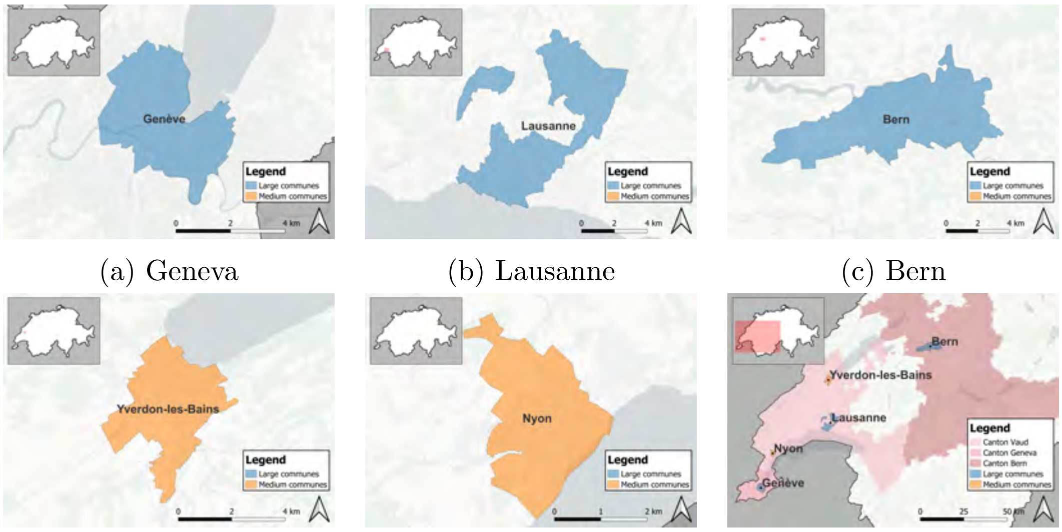


Figure 1-2: Car use vs. population and population density of the major Swiss agglomerations. Data from OFS and ARE (2017) [3].

However, at this point, it is unclear whether these differences could be explained by morphological/geographic differences in Lausanne relative to the other agglomerations. We will thus enter, in the next section, a brief comparative analysis of similar cities to tease out the particularities of Lausanne that might additionally help explain these differences. It should once again be noted that this comparison will be purely for the purposes of establishing the context; the rest of this work will take the form of a case study centered on Lausanne.

1.2.1 Comparison with other Swiss agglomerations

Here we present a brief comparison of Lausanne with five other Swiss urban areas. For comparison's sake, we have selected four cities that are similar geographically and/or in terms of modal practices in an attempt to tease out characteristics unique to Lausanne. In addition–to control for the effect of city size–we pick two similarly-sized agglomerations and two agglomerations with dissimilar size but other comparable characteristics regarding modal practices or geographical situation. In Figure 1-3, below, we present maps of the selected urban areas and a map indicating their spatial relation to one another.



(a) Geneva (b) Lausanne (c) Bern

(d) Yverdon (e) Nyon (f) Overall map

Figure 1-3: Situation of the selected communes in Switzerland.

Table 1-1 below summarizes the population, spatial extent, and population density of each of the selected cities (Geneva, Lausanne, Bern, Yverdon-les-Bains, and Nyon), presented in order of decreasing population. Lausanne lies between Geneva and Bern across all three metrics of population, spatial extent, and population density, positioning Lausanne as an intermediate case among the large agglomerations. Regarding the two medium-sized cities (Yverdon and Nyon), the only immediately apparent similarity is that of population density, by which metric Nyon is the most similar to Lausanne. This is likely due to the large, low-density agricultural area surrounding the higher-density center of Lausanne proper.

City	**Population**	**Area [km^2]**	**Density [hab/km^2]**
Geneva	203,840	15.93	12,796
Lausanne	141,418	41.38	3,418
Bern	134,506	51.62	2,606
Yverdon	29,827	13.52	2,206
Nyon	22,465	6.790	3.309

Table 1-1: Population, area, and density statistics for Lausanne and the four comparison cities. Statistics shown are for communes, not agglomerations. Source: Federal Statistical Office (BFS), 2022.

Topography and situation: Geneva, the largest of the agglomerations, is a strongly multinational city wrapped around the southern tip of Lake Geneva. Surrounded by the Jura to the north and west, Geneva comprises a densely urban center fading into medium-density suburban centers radially from the lake and into France. Lausanne, a historic and large agglomeration in Canton Vaud, rises up above Lake Geneva's northern shore, sprawling across several sloping hills and valleys cut by two small rivers. Bern, Switzerland's political center, is a large agglomeration with a compact historic center split in two by the river Aare, whose urban form is shaped by the river's meandering course. Yverdon-les-Bains, again in Canton Vaud, is a medium urban area nestled in the flat delta at the southern end of Lake Neuchatel. Finally, Nyon, on Lake Geneva's western shore in Canton Vaud, is a medium city with a small historic bourg which quickly gives way to agriculture as the flat shores of the lake melt into the foothills of the Jura mountains.

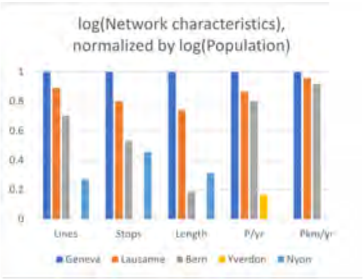
Public transport situation: Here it is also pertinent to discuss the public transport situations in each of the cities. The networks are presented schematically in Figure A-1. The complexity and spatial extent of the networks appear to correspond with city size: Geneva and Bern's are quite complex with a mix of radial and tangential lines; Lausanne's network comprises largely of lines connecting the suburban centers to the west; and Yverdon and Nyon comprise only a handful of lines radiating out from the train station. The networks are described quantitatively in Table 1-2.

City	**Number of lines**	**Number of stops**	**Network len. [km]**	**2022 Usage [M. pass.]**	**2022 Usage [M. pass.-km]**
Geneva	65 bus, 5 tram	1,100⁽¹⁾	750⁽²⁾	227.8	385⁽²⁾
Lausanne	45 bus, 2 metro, 1 regional rail	600⁽¹⁾	375.7	114.2	251.4
Bern	26 bus, 5 tram	343	128	87.4	195
Yverdon	5 bus	65	50.6	5.5	N/A
Nyon	8 bus	125	75⁽¹⁾	3	N/A

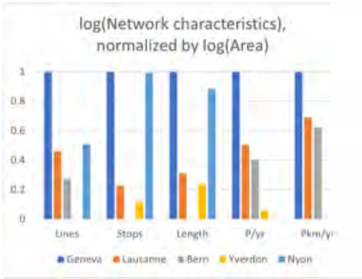
(1) Estimated from network maps. (2) Trams comprise only 7% of the network extent, but account for 50% of the total annual pass-km.

Table 1-2: Summary of public transport networks for the six urban areas, sourced from the respective public transport agencies (tpg, t-l, BERNMOBIL, Travys, and TPN)

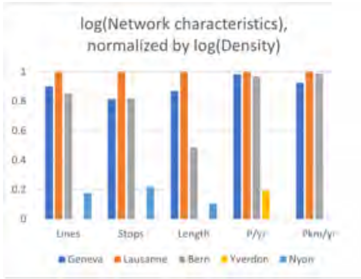
We notice for all cities an apparent directly proportional trend in network intensity characteristics with increasing city size (except for Nyon, which boasts a network almost twice as dense as that of the next-largest city of Yverdon). However, normalizing with respect to population, surface area, and density is much more telling in terms of what actually varies from one city to the next. Figure 1-4 below shows the results of the PT attributes for each city normalized by population, surface area, and population density.



(a) Attributes normalized by population.



(b) Attributes normalized by surface area.

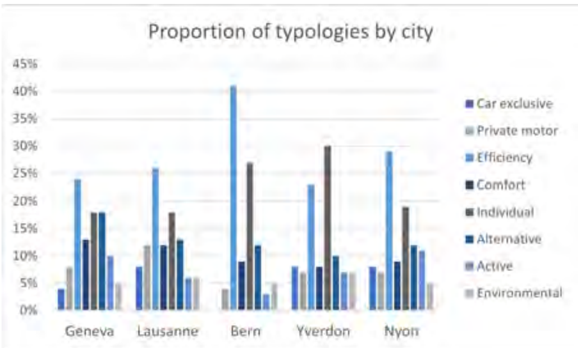


(c) Attributes normalized by population density.

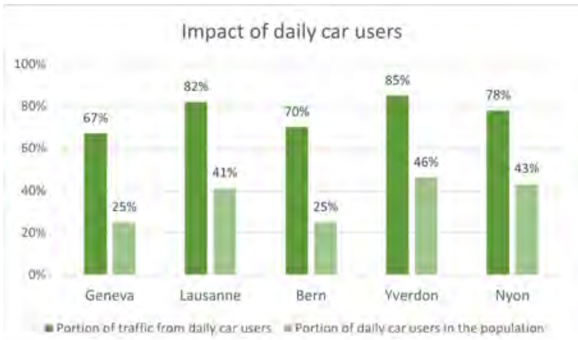
Figure 1-4: Graphs of normalized network characteristics for the five cities.

Lausanne (in orange) is once again consistently between the two other large cities, with a network not quite as extensive as Geneva, yet not as light as Bern. Interestingly, however, Lausanne presents, across all attributes, the highest intensity of public transport network development and use with respect to population density. Nevertheless, normalized by population and surface area, Lausanne remains in-line with the trend.

Modal splits: The discussion in this section is based largely on the work done in an extensive comparative study of mobility practices and habits across Western Switzerland performed by Dr. Kaufmann at the LaSUR lab of EPFL in 2020 [4]. This report segments a survey sample in each city into eight user typologies based on use frequency of active, car, and public transport modes, personal values, and positive/negative perception of these modes. These user typologies attempt to explain individual mode choice through personal values. The categories and descriptive characteristics are presented in Table A-1. From these typologies, the following distributions of mode choice logic arise:



(a) Mode choice logic distribution by city.



(b) Daily car users' impact on urban traffic.

Figure 1-5: Individual mode choice behavior by city, adapted from [4].

Among the large cities, Lausanne has the highest proportion of these car-exclusive and private-motor-predisposed individuals, comparable to the medium cities. Lausanne's share of alternative and active mode-predisposed individuals are similar to the medium cities, while the share of those prioritizing comfort, efficiency, and individual modes are more comparable to Geneva. This suggests a car-reliance in Lausanne comparable to the smaller cities, while the expectations of other modes (in terms of comfort, efficiency, and individuality) are more akin to the larger and denser Geneva. Supporting this notion is Figure 1-5b showing the amount of car traffic attributable to daily users in each of the cities, which reinforces the idea that Lausanne's car users are few but impactful. However, even though daily car users account for 41% of the population, only 8% of the population are "car-exclusive." This implies that at least 33% of the population drive cars daily and are open to other modes, as long as these other modes are efficient, comfortable, or allow for individual space. Hence, targeting this population becomes an incredibly important lever of action for driving the modal shift away from regular automobile use.

This example also emphasizes the divide between rationality and mode choice logic: by all examined metrics and statistics, Lausanne is one of the cities best served by public transport, even accounting for differences in population, area, and density, yet it comprises the highest proportion of individuals with strong (daily) automobile habits. This dichotomy alone demonstrates that there are other forces at play driving individual mode "choice"; if everyone acted perfectly rationally, the decision to take an automobile would gradually disappear with increasing public transport network density — but we do not observe this. Instead, mode choice varies by location independently of size or intensity characteristics, suggesting social aspects at play instead. This is why it is necessary to develop a new schema to begin to diagnose this seemingly irrational, individual mode choice, to break away from this orthodoxy of statistics and adopt a more human-centric approach.

1.2.2 Hypotheses and observations

I present here a series of observations and hypotheses regarding mode choice, dependence, and shift in Lausanne, following the method proposed by LeMieux [8]:

1. Mode choice has a strong socio-spatial component; it varies according to city center vs. periphery, age, gender, income, and even degree of education [5].

2. Car users are particularly reluctant to give up their cars in Lausanne relative to other Swiss agglomerations, despite not having overwhelmingly car-exclusive attitudes [2][3][4].

3. Planning authorities are eager to reduce car use and drive a modal shift toward public transit and have enacted many policies to do so based on traditional planning tools, but car use persists [5][7, p.31-34].

We may thus make the following logical deductions:

1. Individuals do not act perfectly rationally or economically (*homo oeconomicus*), as many planning models and tools assume [7, p.43-46].

2. There must exist some factors not captured by existing planning tools which can explain this individual modal inertia regarding the shift toward public transport [9, p.319-20].

3. These individual dispositions to using public transit manifest themselves in individuals' travel patterns over both long and short time horizons [10, p.298-99]; [5].

4. These dispositions must be extremely influential for informing effective planning decisions and driving the modal shift towards public transport [11, p.125-26].

The central question of this research is therefore to understand how the gaps between observed travel behavior and current planning tools in the Lausanne agglomeration offer insight to individual attractors and resistances to the modal shift from cars toward public transit, in order to better inform future planning tools and decisions. We first perform a review of the literature, to establish the state of the art of such studies, as well as to establish a common understanding of central terms and concepts.

1.3 Literature review

In this section we present a cursory review of the current existing literature surrounding mode choice habits and their quantification. We begin by establishing a conceptual framework from current literature, before discussing methods of quantification, their role in transportation planning tools, and how habits relate to network service gaps.

1.3.1 Habits: a conceptual framework

The concept of habits surrounding mode choice is particularly complex to understand exactly, due to both varying conceptual and operational definitions. Aristotle is widely credited with having laid the groundwork for our conception of habit, with his notions of *ethos* — an active disposition in which individuals are receptive to new schemas — which settles by repetition and reinforcement into *hexis* — an internalized tendency of thought and action which is socially transferrable [12]. Since then, numerous disciplines have weighed in with their interpretations and frameworks, most notably in the social and economic sciences, regarding the specific relationship between habits and behavior. For example, economists such as Gary Becker [13] tend to interpret habits as no more than a repetition of perfectly rational, deliberate choices that are serially correlated, while at the other end of the spectrum, sociologist Bernard Lahire (1998) posits that not only are habits not purely physiological, but can in fact be just as much "socially constructed, in repetition and formal or informal reinforcement" [14, p.89, own translation].

In 2010, Geoffrey M. Hodgson developed a framework of habits against these mainstream *homo oeconomicus* conceptions, building on Darwinism and evidence from brain imaging studies to assert that the relationship between habit and choice is in fact somewhere between fully rational and deliberate and fully subconscious. For instance, in individual human development, instinct manifests before habit, which in turn precedes belief, which subsequently precedes reason. Fundamentally, this means that, according to principles of Darwinism:

> The capacity for belief and reason develops on a foundation of acquired instinctive and habitual dispositions... [Instinct, habit, belief, and reason] are arranged in a hierarchy of functional dependence, where the current operation of reason depends upon belief, belief depends upon habit, and habit depends upon instinct. Lower elements in the hierarchy do not entirely determine the higher functions, but they impel them into being, where they are formed in their respective natural and social context. [15, p. 7].

In this light, Hodgson views the formation of habits as an evolutionary advantage to save computational expenditure on the energy-intensive process of total rationality, whereby experiences are internalized into habits and integrated, along with a reduced rationality demand, into future decisions. This conception is remarkably similar to Aristotle's *ethos* and *hexis*, and is even validated by brain scans which show the process of habit formation involving a shift of brain activity from the area of the brain responsible for conscious decision-making (the pre-frontal cortex) toward that associated with context-triggered responses (the basal ganglia) [16]. Hodgson's central assertion — in-line with other Darwinian psychologists such as William James, Thorstein Veblen, and John Dewey — is that habits are a semi-conscious, socially-formed and transmitted, necessary but not sufficient, context-specific causal mechanism on human behavior that describes a "**propensity to behave in a particular way in a particular class of situations**" [15, p. 4].

Social psychologist Bas Verplanken developed a similar framework from the perspective of the social sciences, with the intent to operationalize this conception into an index of habit strength (discussed further in Section 1.3.2). Verplanken anchors this definition around five core aspects of habits, namely: repetition, effortlessness, automaticity, efficiency, and identity [17]. Acknowledging that habitual behaviors are formed by "frequently and satisfactorily pairing the execution of an act in response to a specific cue," [17, p. 1314], Verplanken also notes that repetition alone does not guarantee the formation of a habit, but simply increases its probability, which, by extension, means that the activation frequency of the behavior is entirely separate from the strength of the habit.

In his work *Déplacements urbains: sortir de l'orthodoxie : Plaidoyer pour une prise en compte des habitudes* (in English: *Urban movements: emerge from the orthodoxy: plea for a consideration of habits*) [7], Thomas Buhler extends Verplanken's discussion of this common conflation between frequency and choice in order to provide an even more operational framework as it pertains specifically to mode choice. Buhler once more asserts that behavioral repetition and regularity, although crucial in the habit formation stage, do not determine the strength of a habit: regardless of the strength of a habit, once it is formed in an initial context, the behavior may remain dormant until the initial context reoccurs. Thus, the behavioral activation frequency *after* the formation phase is entirely independent of the behavioral activation frequency *during* the formation phase, which means the overall frequency of the behavioral activations cannot be used to measure the habit strength. Instead, the behavior can only be activated within the initial context, invoking Hodgson's notion of a propensity for activation in a given context; habit strength is not causally linked to the behavior's overall activation frequency. This conflation of *mode use frequency* with *mode use habit* nevertheless persists in the habit quantification literature, as Hodgson underlined earlier with the work of economist Gary Becker [13].

Centered in this definition of habits, Buhler identifies two principal axes by which we can diagnose the strength of a habit according to context: an axis of "intensity," which describes the frequency at which a behavior is activated in a singular class of situations (i.e. going shopping, going to work); and an axis of "breadth," which describes the variety of different classes of situations in which a behavior is activated [7]. This description is presented graphically below.

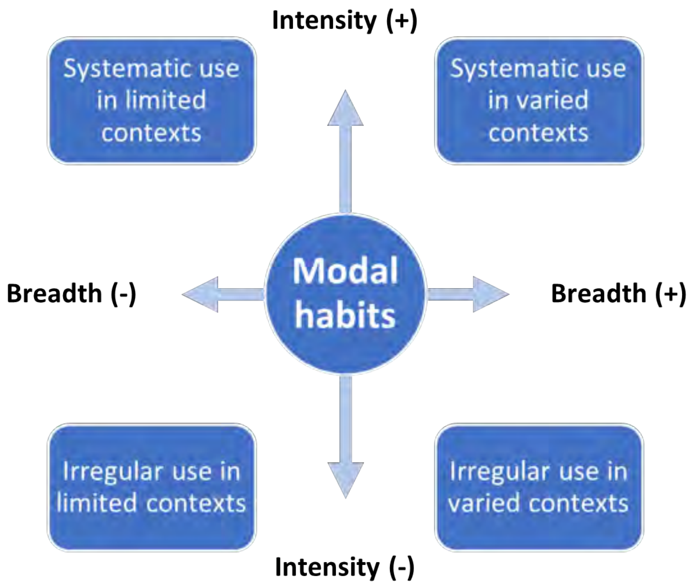


Figure 1-6: The two dimensions of habit strength in the context of mode choice (after Buhler, 2015) [7].

For example, an individual who regularly takes their car to go grocery shopping would be said to have a high-intensity habit (in the context of grocery shopping). If this person also took their car regularly to go to work and visit with friends, they would be said to have a high-breadth habit, as well. This framework proposed by Buhler offers a useful anchor point for us to quantify habits.

1.3.2 Quantifying habits

Quantifying habits remains a black box within the literature. The best-established method for quantifying the strength of a habit is the SRHI, or, the Self-Reported Habit Index proposed by Bas Verplanken et al. [17]. Anchored to five core aspects of habits, the SRHI takes the form of a 12-question quiz, with the responses scored on a five-point Likert scale. Example questions include: "it's something I do without thinking", or "it's something I've done for a long time". From these questions, it is clear to see that the definition of "habit" baked into the formulation of the questions considers habits more like a reflex than a propensity for a behavior to be activated in a given context. In fact, the SRHI intentionally avoids investigating habit cues or behavioral contexts, because it aims to be a general instrument, and, as such, only provides a single-number score. Because of this, we depart from this tool in favor of an approach that infers modal habits through *observed* travel behavior–despite additional challenges–with an explicit aim to link habit cues (daily activities) with habitual behavior (mode choice). We present the proposed methodology to address this problem in Section 2.3.1. Nevertheless, we will validate the method against the SRHI, which remains the best-established method of habit quantification [7].

It is also common to focus on "time budgets" and sequence analysis in habit characterization [18][19][20], where the 24-hour day is discretized into time windows, and an individual's mode choice or activity ("state") is recorded in each window. This is helpful for a sequence analysis approach to habit quantification, particularly where researchers intend to quantify how much an individual's habits change from one day to the next. In the case of [18], which used a Hidden Markov Model (HMM) to characterize these sequences, the mode choice in each discrete window (deemed the "observed state") was considered to be determined entirely by the location (deemed "hidden state"), which in turn is determined exclusively by the previous location choice. This approach directly integrates sequential dependencies between successive location and mode choices, as well as the joint dependency between location and mode choice. This ultimately allows for a combined assessment of space, time, and purpose, all of which play a significant role in the conceptualization of habits as being highly context-specific.

1.3.3 Current planning tools

In this paper, we refer to "planning tools" to mean quantitative methods or metrics that are used to inform planning and decision-making. This includes methods ranging in complexity from simple non-spatial indicators, to composite spatial indices, to full numerical mode-choice models. These are the three broad categories of tools that will be examined in this work, and a more thorough description of the current state-of-the-art for each of these categories is presented in the following paragraphs.

I. Service indicators are generally single numbers that represent an average attribute of the public transport service. Eboli [21] presents a comprehensive review of common and novel objective public transport performance quality indicators across nine service aspects: service availability, service reliability, comfort, cleanliness, safety, fares, information, customer care, and environmental impact. A list of indicators reviewed by Eboli is presented in the Appendix, in Table A-2. This planning tool will not be explored in-depth, but rather will be used in the mode choice model to evaluate PT alternatives. It remains one of the most easily-interpretable ways to persuade PT operators to target network improvements.[1]

II. Composite spatial indices are perhaps the least rigorously-defined of the three. Because each public authority, planning agency, and public transport operator will have different values and priorities (both because of cultural/geographic differences and structural/institutional/operational differences), each of these organizations are likely to develop their own composite spatial indices of network quality based on what is important to them. For this reason, we examine the spatial index used by the Canton of Vaud to assess network quality. The cantonal index integrates hierarchical preference across three different dimensions: modes of transport (i.e. train, tram, bus); frequency (time interval between vehicles, for all modes of transport); and accessibility (straight-line distance from the stop/station). Ultimately, each point in space is assigned to a service quality category ranging from A (excellent quality) to E (low quality), which increases with proximity to public transit stops and intensity of service (high frequency and/or high capacity according to mode) [22]. This approach makes sense on a national or even regional scale, but for more fine-grained analysis, there is a conflation of service intensity

[1]This is also corroborated by Redman (2010) [11], where, in a meta-review of various service quality evaluation schemes, reliability was consistently found to be especially important in persuading car users.

and the service connectivity – as, at smaller scales, the national and regional networks are less important for intra-city mobility.

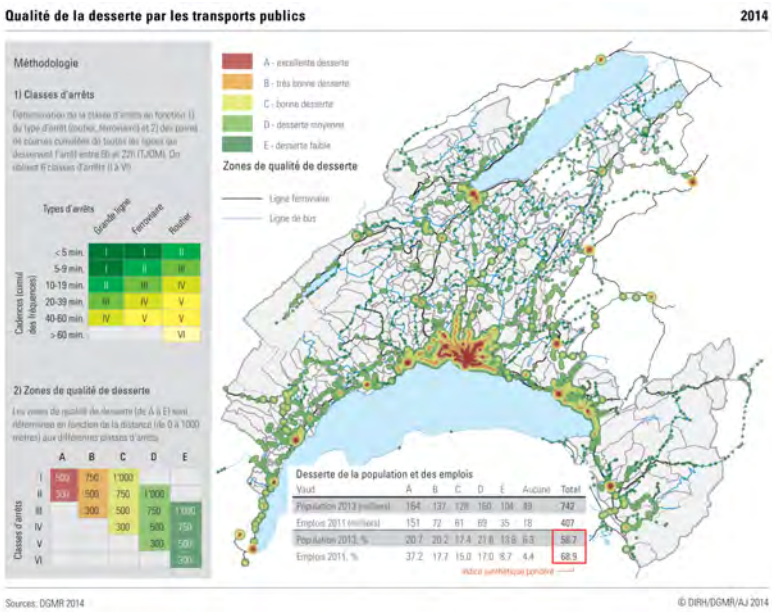


Figure 1-7: Illustration of the Vaudois service quality index, from [22, p.14].

III. Numerical mode choice models are the most complex category of planning tool considered here, and are most commonly logit models. This includes the Swiss National Mode Choice Model (NPVM), which was recently established and calibrated in line with the 2015 mobility micro-census. The purpose of the model's establishment was to understand the movement of people between Swiss cities according to transportation mode. Consequently, the model takes the form of a nested logit model, with the first decision being the travel mode and the second being the destination. This, of course, is predicated on the assumption that the choice of destination depends on the mode choice. There are thus three "nests" (one for each of the investigated transportation modes: car, public transport, and "light" modes), and 11 destinations within the nests, corresponding to 11 major Swiss agglomerations [23]. The model was calibrated on a national basis, considering mobility between cities according to the results of the 2015 micro-census.

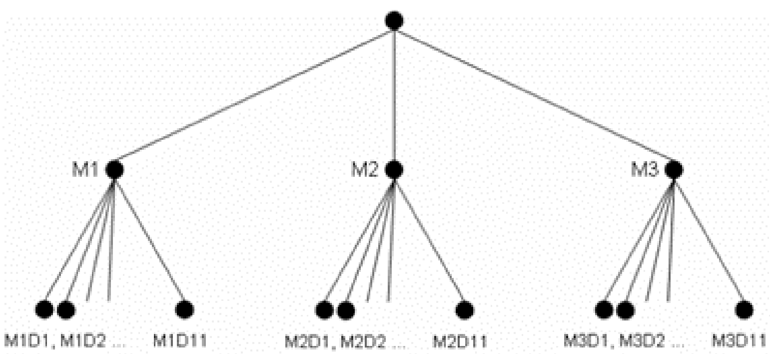


Figure 1-8: Graphical representation of the NPVM structure, from [23].

The utility equations used to determine relative mode choice probabilities are presented below.

$$
\begin{aligned}
U_{\text{car}} &= ASC_{\text{car}} + \text{time}_{\text{car}} \cdot \beta_{\text{car,time}} + \text{cost}_{\text{car}} \cdot \beta_{\text{car,cost}} + \text{availability}_{\text{car}} \cdot \beta_{\text{car,avail}} \\
&\quad + \text{employed} \cdot \beta_{\text{employed}} + \text{work parking} \cdot \beta_{\text{work parking}} \qquad (1) \\
U_{\text{PT}} &= \text{time}_{\text{PT}} \cdot \beta_{\text{PT,time}} + \text{cost}_{\text{PT}} \cdot \beta_{\text{PT,cost}} + \text{access time} \cdot \beta_{\text{access}} + \text{transfers} \cdot \beta_{\text{transfers}} \\
&\quad + \text{frequency} \cdot \beta_{\text{frequency}} + \text{age} \cdot \beta_{\text{age}} + \text{GA} \cdot \beta_{\text{GA}} + \text{demitarif} \cdot \beta_{\text{demitarif}} \\
&\quad + \text{employed} \cdot \beta_{\text{employed}} + \text{work parking} \cdot \beta_{\text{work parking}} \qquad (2) \\
U_{\text{active}} &= ASC_{\text{active}} + \text{time}_{\text{active}} \cdot \beta_{\text{active,time}} \\
&\quad + \text{employed} \cdot \beta_{\text{employed}} + \text{work parking} \cdot \beta_{\text{work parking}} \qquad (3)
\end{aligned}
$$

1.3.4 Behavior/planning gaps

Various authors have criticized the metrics used by public transport suppliers to assess network service quality. Rietveld [9] argues that supply-oriented metrics are fundamentally misaligned with individuals' perception of the experience. While suppliers are prioritize aggregate level metrics (i.e., average frequency, average expected wait times, etc.), user perceptions are shaped by extreme experiences (i.e. even a single 10-minute delay). Friman [24] assigns the term "critical incidents" to such isolated extreme instances, which–whether positive like a friendly bus driver, or negative as in the case of an extreme delay–are pivotal in forming user perception, but are historically rarely considered in the supplier's evaluations of the service quality. In working directly with a public transport supplier, Parkan [25] remarked that the supplier explicitly didn't consider the users' perception of service quality in their evaluation of network performance and were instead almost overly concerned with the optimal balance of economic ratios and targets. The same study noted that in fact much of the contemporary research on supplier performance metrics is based around such economic ratios with no regard for the individuals' interactions with the physical network–perhaps due to the cost of quantitatively evaluating such interactions, normally done through surveys.

There also exist numerous studies corroborating the gaps between suppliers' indicators of service quality in common use and the actual way the service is used and perceived individually. Rietveld (2005) makes the argument that this is due to fundamental, structural differences between what is important in operations versus use [9]. In a more specific example, Chien [26] found that the optimal spacing of bus stops depends strongly on the individual value of time; thus, public transport accessibility is not objective but rather heterogeneous within a population and cannot appropriately be approximated with a single rule-of-thumb number. El-Geneidy [27] confirmed this in 2014, finding that, for one urban area, the rule-of-thumb minimum service area of 400 meters didn't align at all with the actual radius within which people used public transport services, due to the aforementioned heterogeneity of perceived accessibility.

To quantify these gaps between planning and individual behavior, the literature contains multiple different approaches. One common approach begins with a standard supply/demand gap analysis, introducing different layers or variations such as computing the gaps through time and space [28]; [29], or even across user typologies [30]. Schultheiss' recent approach investigated the supply-demand gap for public transport in Geneva over the course of the day, as well as according to user typology [30]. Here, eight user typologies were established via hierarchical clustering in the same manner established by Drevon and Gumy [31]. This provides a much more comprehensive understanding of specific deficiencies in the network and can provide insights into the types of improvements to make in order to target certain segments of the population.

1.4 Research question and positioning

The research aims to unlock the role of individual habits in mode choice and attempt to derive relationships between these habits and current planning tools in order to better inform future planning decisions. As such, the research will concentrate on two of Dupuy's [32] four "urban networks": 1. individuals (whose behaviors are informed by habits) and 2. public transport operators (whose behaviors are informed by planning tools). These two agents interact with one another through the public transport network: individuals through mode choice, and public transport suppliers through network improvements. Each perceive the network as well as behavior of the other in different ways, which they use to inform (consciously or unconsciously) their future interactions with the network. Figure 1-9 below summarizes these key actors, their behavior, and their behavioral influences involved in the framework of the research question.

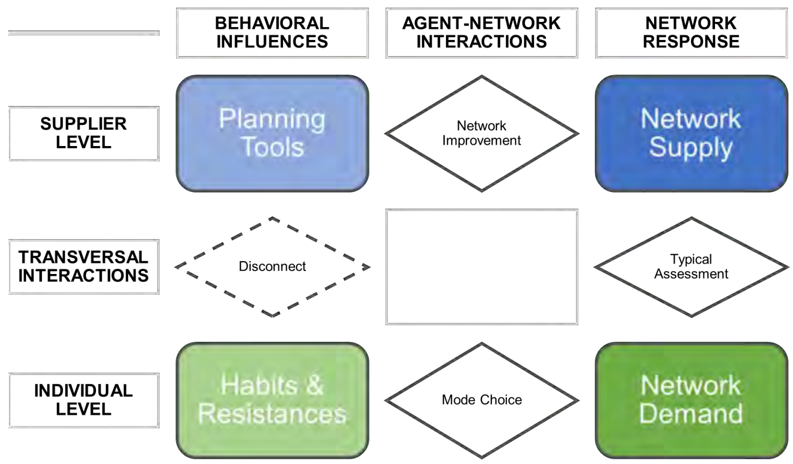


Figure 1-9: Diagram of actor behavior and influences on behavior through interaction with the urban public transport network.

State-of-the-art assessments of network service quality, as discussed above, tend to focus on the interaction between network supply and demand. However, it has also been documented that individual habits and resistances play a major role in informing mode choice and, ultimately, network demand. When network operators ignore this aspect, it becomes incredibly difficult to provide a service that accurately caters to the population being served. This is because rather than explicitly considering behavioral motivations, network operators must instead consider the effects of these behaviors, filtered through multiple levels of interaction with the system. Consequently, persuading a holdout group of strongly car-dependent individuals toward the public transport service becomes almost a matter of guesswork. This research thus attempts to close the gap between planning tools used in informing network improvements, and individual behavioral logic—i.e. habits and resistances—informing modal shift dynamics. As an outcome, all three categories of existing planning tools will be evaluated, and new planning tools (spatial indices and mode choice models) will be proposed to help public transport operators understand and quantify the spatial distribution of these individual habits and resistances on a population-wide scale, as well as make more targeted network improvements to kick-start the modal shift for both highly car-dependent users and other users resistant to using public transport.

2 Materials and Methods

2.1 Study perimeter

Figure 2-1 below depicts the perimeter of study, which shall be based on the definition of the Lausanne-Morges agglomeration according to the Projet d'Agglomération Lausanne-Morges (PALM) [33], the spatial extent at which public transport and other mobility infrastructure improvements are planned for the agglomeration. The agglomeration covers over 60 square kilometers, 26 communes, and approximately 294,000 inhabitants.

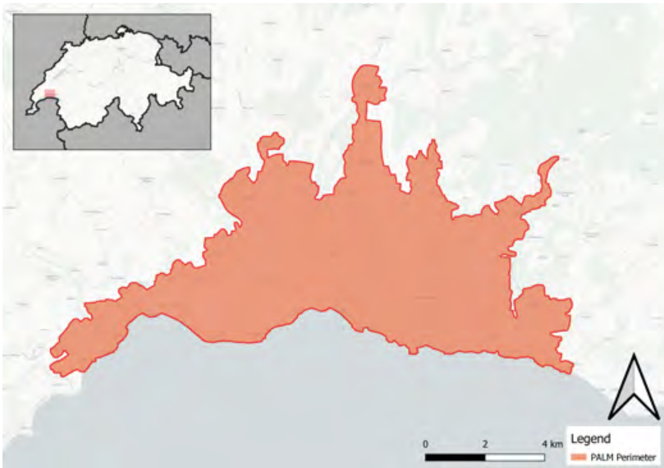


Figure 2-1: Perimeter of study, according to the perimeter of the PALM (adapted from État de Vaud, 2023).

2.2 Lemanic Panel data

The data used for this work comes from the Lemanic survey panel conducted between April 24, 2023 and June 5, 2023 in Switzerland from the Laboratory of Urban Sociology (LaSUR) at EPFL. The panel comprised of survey and GPS tracking data of 2,802 individuals over the course of several weeks. The average observation period was 36 days, resulting in a total of 668,242 activities for users based in the Lake Geneva region of French-speaking Switzerland and France (see Figure 2-2). The data was then split into staypoints and legs, each containing attributes like purpose (labeled by the individuals), start time, end time, and inferred travel mode (inferred afterwards from GPS data). The data was available pre-filtered spatially and temporally to remove unusable data, with a loss of only as much as 5%.

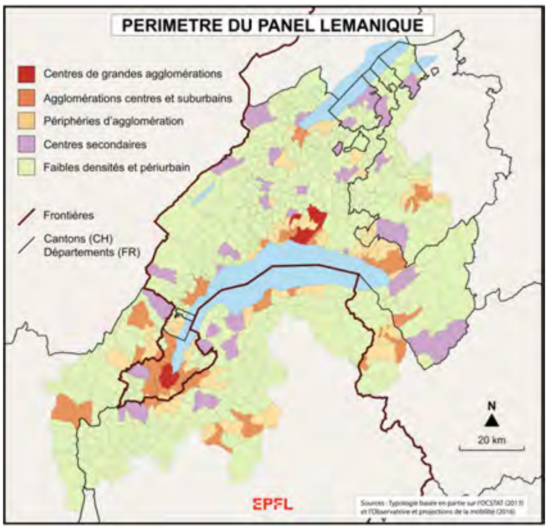


Figure 2-2: Perimeter and urban typology of the Lemanic Panel sample.

The data was subsequently filtered again to extract all individuals with at least one journey with a 'home' purpose ending within the study perimeter (Figure 2-1). This resulted in a final sample of 767 individuals with 201,101 trips in total. An individual's 'home' location was inferred as the mean coordinates of the endpoints of all trips with a 'home' purpose within 200 meters of one another, which allows for the preservation of multiple home locations, in the case of multiple residences. This 200m buffer is in line with [10], who identified this distance as the farthest away from their home that people normally park their car.

2.3 Analytical framework

The analytical framework comprises three main steps: 1. quantify individual habits; 2. identify spatial gaps in network supply corresponding to the spatial distribution of modal habits; and, 3. calibrate and propose a numerical mode choice model to Lausanne (see Figure 2-3). The subsequent section elaborates the approach to each of these three steps.

Figure 2-3: Analytical framework.

2.3.1 Habit quantification

The framework used to quantify habits in this paper will be strongly anchored in Buhler's axes of intensity and breadth presented earlier in Figure 1-6. Habits are inscribed in a network of habitual places, purposes, and scripts, wherein, depending on the strength of the habit, the choice a successive activity falls somewhere on the spectrum of being strongly connected with other habitual places, or may be almost entirely spontaneous and improvised [7, p.95-96]. Through this lens, the strongest indicators of modal habits will thus be these habitual places and purposes, which can serve to identify the network of modal habits in which other behaviors might be inscribed, or with a potential to be inscribed. For example, a commuter taking the same route every day may start to notice new shops or cafés along the route and one day spontaneously decide to pop in and take a look around. Subsequently, if this new place is attractive enough, the commuter may decide periodically to repeat this behavior, and over time, this detour may itself develop into a habit.

Given the strong dependence of mode choice on destination and purpose, as well as the dependence of mode choice on the previous choice, we propose to use a Hidden Markov Model (HMM) approach to explain the sequence of mode choices using the sequence of activities. Through numerical simulation, we extrapolate the observed travel behavior to equilibrium and directly observe these conditional probabilities for the calculation of habit intensity and breadth of transportation mode, considering structural interdependencies between modes, between locations, and between modes and locations, similarly to [18].

Hidden Markov Model (HMM) is an extended version of a classic Markov chain model which differs in several key ways. Where a standard first-order Markov model (based on the fundamental assumption that the action at time t depends only on the action at time $t-1$) explains a sequence of actions based only on the probability of transitioning between each pair of possible actions, the Hidden Markov Model explains the probability of an action based on a non-observed, i.e. "hidden," Markov variable z which fully determines the choice x at time t. Figure 2-4 below demonstrates these relationships schematically.

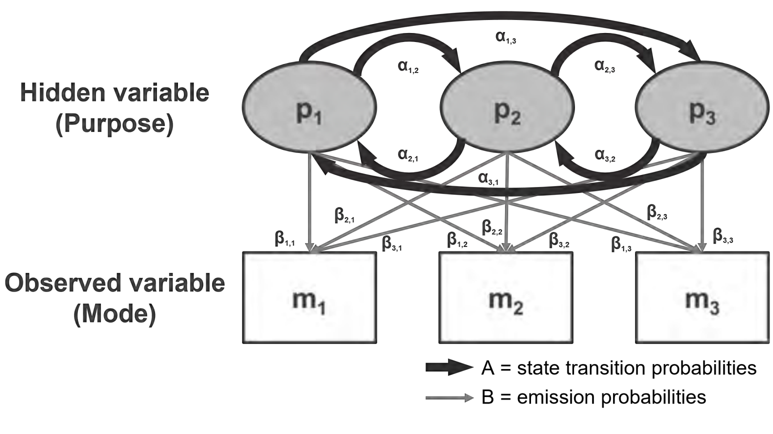


Figure 2-4: Diagram of an example HMM model for mode choice.

The fundamental assumption is then $P(x_t) = P(x_t|z_t)$. In other words, the occurrence of item x_t in the sequence of observed values x (in this case representing the sequence of mode choices) is determined by the corresponding item z_t in the sequence of hidden states

z (in this case representing the sequence of trip categories). An HMM then estimates the following parameters: the *equilibrium vector* π of probabilities of being in each observed state at equilibrium; the square *transition matrix* A of probabilities of transitioning between hidden states i, j at equilibrium; and the *emission matrix* B of size (number of hidden states, number of observed states) storing the conditional equilibrium "emission" probabilities of observing action x_j given hidden state z_i.

One other key difference between HMM and standard Markov models is that the equilibrium probabilities for standard models may be resolved as the eigenvector of a transition matrix, while these probabilities must be estimated through numeric simulation for HMMs. For a more thorough discussion of the validity of the use of HMM to mode choice, see Appendix A.3.

In order to quantify an individual's habits, an HMM is fit to each individual through numeric simulation, with initial transition probabilities estimated from the mode and purpose sequences.[2] The simulation is repeated for a predetermined number of iterations, then averaged over all runs. Simulations were initially run for 1,000 iterations, but the variance of these initial samples was found to converge to within 5% of this final value after only $n = 50$ iterations. The trip attribute used for habit quantification was chosen to be location "importance", which we define as the average number of weekly visits to a geographic location (a visit is considered repeated if the endpoint is within 200m Euclidean distance of a previous endpoint). All sequences must comprise categorical variables: sequences of strings (modes and purposes) are grouped and encoded as unique integers, while sequences of continuous variables (distance and time) are split into discrete bins, where the bins are then encoded as unique integers. The thresholds for binning importance values are: 0.5, 1.0, 3.5, 7.0, and 14.0. Table 2-1 summarizes the grouping of modes and trip purposes.

	ID	Group name	Labels in group
Mode	0	Car	"Car", "E-car", "Carsharing", "Taxi/Uber"
	1	Public transport	"Bus", "Train", "Tram", "Subway"
	2	Active	"Bicycle", "Kick scooter", "Bike sharing"
	3	Alternatives	"Motorbike", "E-Bike"
	4	Other	"Airplane", "Boat", "Other"
	5	Walk	"Walk"
Purpose	0	Home	"home"
	1	Family/ friends	"family/friends"
	2	Work/ school	"work", "study"
	3	Errands	"errand", "shop", "medical visit", "assistance"
	4	Free time	"eat", "leisure", "sport"
	5	Other	"wait", "unknown", "other"

Table 2-1: Grouping of categorical variables for trip mode and trip purpose.

Following these encodings, Figure 2-5 presents a sample encoding procedure:

[2]This can also be done using Maximum Likelihood Estimation if the sequences of the hidden state and observed state are both known, as in this case. However, due to time constraints, this could not be implemented and remains a step for future research.

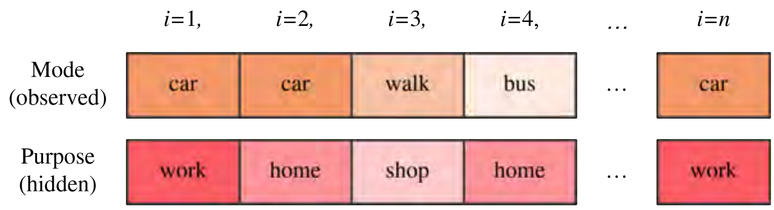


Figure 2-5: Sample vectorization of mode and activity chains.

It is interesting to note that two locations may be geographically distinct, but share a purpose (i.e. different shopping centers), or fall into the same distance class. In this way, it could be interesting to compare habit strength results with different hidden state schemes. However, trip importance was ultimately selected to be the segmenting characteristic. Ideally, trips would have been segmented by purpose, but due to the data being self-labeled by the survey participants, approximately 40% of the trips were lacking a labeled purpose. Model fitting was initially repeated for various habitual contexts: trip purpose, geographic location, trip distance, trip duration and trip "importance." The results were similar enough between categories that ultimately trip importance was chosen as a label-agnostic proxy for trip purpose.

Intensity is defined by Buhler as the "frequency at which a specific behavior is activated in a given category of situations" [7, p.68, own translation]. From this definition, there arise three key terms to define: "frequency of activation," "specific behavior," and "category of situations." "Specific behavior" is perhaps the easiest of the three to define: in our case of transportation mode habits, the specific behavior is simply the mode of interest for a given category of situations. The "frequency of activation," then, is the frequency that this mode is used for a given category of situations. Finally, the "category of situations" remains trickier to unpack. For clarity, we replace the term *category of situations* with *habitual context*, within which there are discrete categories. For example, a habitual context could be trip purpose or location importance, which could contain categories of (home, work, shop, etc.) or (locations with $1.0 - 3.5$ weekly visits, locations with $3.5 - 7.0$ weekly visits etc.), respectively. As previously mentioned, we will focus only on location importance, but further work could compare different trip attributes.

Given that the intensity of a habit is specific to a category of situations, a single intensity value must be derived for each situational category, then aggregated to determine an overall intensity value for the given habitual context. This can be considered as the conditional choice probability $P(m|c)$ of choosing mode m given category c, which exactly describes the entries in the emission matrix B estimated directly from the HMM mentioned above. In order to aggregate these percentages to obtain a single intensity value for a given habitual context, however, we should not use the simple arithmetic mean. This would assume that all trip purposes are of equal importance, which is not the case. In fact, trips to locations with higher visits should be considered more important to a habit, as we have seen that these places reinforce an individual's spatio-cognitive mode-habit network and can serve as catalysts for developing further habits [7, p.96]. The quadratic mean $M(x_1, ..., x_n) = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}$ fulfills this condition by pulling the arithmetic mean toward contexts with higher intensity percentages. Substituting our relevant values into this equation, we obtain the following formulation for the aggregate intensity I of individual n for mode m and habitual context C :

$$I_{n,m,C} = \sqrt{\frac{\sum_{c \in C} P_n(m|c)^2}{n_C}} \in [0,1] \tag{4}$$

where $P_n(m|c)^2$ is the conditional probability of individual n using mode m for category c, estimated from the HMM, and n_C is the number of categories in the habitual context C. Values of I close to 1 indicate a high probability of choosing a particular mode across trip categories, while values of I close to 0 indicate a low propensity for choosing the mode across trip categories.

Breadth is defined by Buhler as the "variety of situational categories in which a behavior is activated" [7, p.68, own translation]. In other words, high breadth means the habit pervades all situational contexts of a person's life, and they are equally likely to choose the given mode for any of the purposes in their choice set. The breadth of a habit can also be calculated from the results of the HMM estimation, because this notion of habit breadth is in line with an individual's aggregate propensity to choose a destination for a given mode. Thus, we derive the individual's conditional choice probability $P(c|m)$ to choose category c given mode m using Bayes' theorem $P(c|m) = \frac{P(m|c) \cdot P(c)}{\sum_{c \in C} P(m|c) \cdot P(c)}$ which yields the conditional choice probabilities for each category c given each mode m, where $P(c)$ is the HMM-estimated equilibrium state probability π_c of category c.

To evaluate the variety, or heterogeneity, of these probabilities, we use Shannon entropy over the entire set of categories C, normalized to the maximum possible Shannon entropy for the given number of categories n_C. Shannon entropy $H = -\sum_i p_i \cdot \log_2(p_i) \in [0,1]$ describes the heterogeneity of a set $\{x_1, x_2, ..., x_n\}$; it has a minimum value of 0 when the set is fully heterogeneous, and higher values indicate greater homogeneity. The entropy of a set of conditional choice probabilities $\{P(c_1|m), P(c_2|m), ..., P(c_{n_C}|m)\}$ indicates the heterogeneity of destination choice for a particular mode: higher entropy for a given mode indicates higher destination choice homogeneity (i.e. the mode is used equally for all trip categories), and lower entropy for a given mode indicates lower choice homogeneity (i.e. the mode is valued more for certain trip categories). The Shannon entropy is first normalized by the maximum possible entropy $H_{max} = \log_2(n_C)$ to compare across individuals with different choice sets, as well as to limit the upper bound to 1. This normalized entropy is subsequently squared to widen the distribution and linearly transformed (keeping the midpoint at 0.5) with a scale parameter α, as observing the maximum and minimum entropies is not likely. We set $\alpha = 2.0$ to stretch the minimum and maximum observed breadth values to 0 and 1, respectively (up until this point, $0.25 \lessapprox B \lessapprox 0.75$ for individuals with ≥ 1 trip). Equation 5 shows the final calculation of habit breadth:

$$B_{n,m,C} = \alpha \left[\left(\frac{-\sum_{c \in C} P(c|m) \cdot \log_2 P(c|m)}{\log_2(n_C)} \right)^2 - 0.5 \right] + 0.5 \in [0,1] \tag{5}$$

where $B_{n,m,C}$ is the breadth of individual n taking mode m in habitual context C, and where $P(c|m) = \frac{P(m|c) \cdot P(c)}{\sum_{c \in C} P(m|c) \cdot P(c)}$ is the conditional choice probability of trip category c given mode m, calculated from the HMM emission matrix. Values of B close to 1 indicate the modal habit pervades all trip categories equally, while values of B close to 0 indicate the habit is activated only in a single context, if at all.

Overall habit strength Q for individual n taking mode m in habitual context C is taken as the simple sum of intensity and breadth:

$$Q_{n,m,C} = I_{n,m,C} + B_{n,m,C}; \in [0, 2] \tag{6}$$

where a mode with a habit of 0.0 is nonexistent, 2.0 is the maximum. These three aspects of habit (breadth, intensity, and total strength) will be used to establish user typologies in further analysis.

Validation against SRHI will be performed to confirm the metric's ability to capture habitual behavior. The Self-Reported Habit Index (SRHI) [17] is, to date, the most well-established measure of habit strength, and thus serves as an appropriate benchmark for our metric[3]. Aspects of the SRHI that had corresponding questions in the Panel data and used for validation were: frequency ("how often do you take this mode?"); flexibility ("in case of disruptions, I am able to change my route or mode without problems"); automaticity ("thinking about the trips you make regularly in your region, to what degree is this mode appropriate for the following types of travel?"); and, for PT only, comfort ("what is your level of comfort with using public transport to get around?"). Two linear regression models were fit to the habit strengths for the motorized modes, which had the most questions in the Panel that corresponded to the SRHI questions: one regression was predicted from 25 socioeconomic variables (null model), and the other regression was predicted from the 3-4 SRHI aspects. R^2 coefficients of the linear-regression habit strength against the actual habit strengths for comparison of the two predictor sets.

User typologies are established using Hierarchical Clustering, similarly to [31], using intensity and breadth for each mode as features.

2.3.2 Behavior/planning gaps

Here the results of the habit calculation are presented spatially, and corresponded with the Vaudois index of PT service quality.

2.3.3 Mode choice model

Modeling assumptions with such a mode choice model are fundamentally that individuals have full knowledge of every aspect of every alternative available to them, and make their choice fully rationally. Although this approach has limited causal explanatory power regarding how individuals internalize the information that is actually available to them on a daily basis, it is nevertheless useful because it offers insight into statistical correlations which can offer a starting point for public policy. We intend to build on this approach with an exploration into the use of habit typologies for enriching such analyses.

As is often customary, travel time and cost are assumed to be perceived differently according to travel mode. Perception of travel cost was assumed to be strongly influenced by income, and was thus transformed in a manner similar to [34] : $\beta_{cost} \cdot$

[3]The full SRHI questionnaire is reported in Appendix, Table A-3.

$(\frac{\text{income}}{\text{mean income}}) \cdot$ travel cost. Perception of overall travel time was assumed to be nonlinear (an additional minute of travel time is perceived differently for short trips and for long trips) and heterogeneous across modes, so the λ parameter of a Box-Cox transformation $(y(\lambda) = \{\frac{y^\lambda - 1}{\lambda}$ if $\lambda \neq 0; \log(y)$ if $\lambda = 0\})$ was estimated for each mode to determine the strength and direction of nonlinearity. Time in public transport was decomposed into five components (access time, waiting time, in-vehicle time, transfer time, and egress time), as each component is likely to be perceived differently.

Subsample selection was necessary due to computational restrictions limiting the maximum number of observations to 7,000 trips, approximately 10 per individual, or 5% of the original sample size. The subsample was generated randomly, in the following manner: iterate through the list of individuals and select, at random, one trip to be moved from the population to the subsample, then move to the next individual and randomly select one trip to move into the subsample, repeating until the subsample is the desired length. This approach ensures that each individual is approximately equally represented in the subsample.

Alternative generation is done to expand the choice set for each individual, in order to compare the aspects of the mode actually chosen with aspects of all available modes. The start and end locations are set by the actual GPS start and end locations, and the departure time was set to the actual departure time. For bike and walking modes, an A* shortest path (by length) was computed for each trip, using the OpenStreetMap (OSM) street network. Additional street aspects available from OSM the Ville de Lausanne (availability of cycling infrastructure, public parking locations) are also used to enrich the feature set.[4] For car and PT modes, the Google Maps API Routing engine was used to calculate optimal public transit connections, car journeys, and their associated details while integrating predicted delays due to traffic at a given departure time. Car fuel costs were calculated based on an assumed average fuel cost of 1.90 CHF/L times an average fuel efficiency of 5.75 L/km [36] ($\approx$ 10.90 CHF/km). Car parking costs were calculated based on the parking duration, then calculated according to an average progressive pricing scheme from the Ville de Lausanne (see Table A-4). Finally, PT ticket costs were calculated according to the actual Mobilis fare schemes for a single ticket [37], based on the number of Mobilis fare zones crossed. Table 2-2 below summarizes the features generated for each mode during this step.

[4]For further consideration of mode choice models specific to bicycle infrastructure, see [35]. From here we draw inspiration for feature selection as it pertains to bicycle route choice, but we depart from this paper by not explicitly exploring route choice.

Category	**Car**	**PT**	**Bike**	**Walk**
Time	Driving time	In-vehicle time Access/egress Transfer/waiting	Cycling time	Walking time
Cost	Fuel cost Parking cost	Ticket cost	–	–
Infrastructure	Work parking	Transfers Frequency	Infra. avail.$^{(1)}$ Ave. slope	Ave. slope

$^{(1)}$ Fraction of route with bike infrastructure, after [35].

Table 2-2: Alternative characteristics generated for use in mode choice model.

Utility equations For final utility specifications implemented in the mode choice model (based on those from the NPVM [23]), see equations A.1–A.4 in Appendix.

Estimation and analysis is done using the open-source Biogeme package for Python[38]. To integrate our notion of habits into this model based on economic theory, we proceed with standard analysis, but segmenting the results by habit typology. The key steps of such a typical analysis are: 1. compute the value of time (VOT); 2. compute direct and cross-elasticities of the mode choice probabilities of each mode relative to a given attribute; and 3. forecast modal splits under different scenarios.

The VOT in linear models is calculated as $\mathrm{VOT} = -\beta_{\mathrm{time}}/\beta_{\mathrm{cost}}$. However, for more complex models with nonlinear variables, the VOT must be determined for each individual by partially deriving the utility function U with respect to both time and cost ($\mathrm{VOT}_i = \frac{\partial U/\partial \mathrm{time}_i}{\partial U/\partial \mathrm{cost}_i}$), then aggregating over the entire population. In cases where travel time is decomposed into τ components, the travel time components are weighted by the value of the travel time itself and normalized by the total travel time, yielding the following equation for individual i's VOT:

$$\mathrm{VOT}_i = \frac{\sum_\tau (\mathrm{time}_{\tau,i} \cdot \frac{\partial U}{\partial \mathrm{time}_{\tau,i}})}{\frac{\partial U}{\partial \mathrm{cost}_i}} \cdot \frac{1}{\sum_\tau \mathrm{time}_{\tau,i}} \tag{7}$$

which is subsequently aggregated according to the statistical weight (socioeconomic representativeness of an individual relative to the communal population) provided in the panel data:

$$\mathrm{VOT} = \frac{\sum_{i=0}^{n} \mathrm{weight}_i \cdot \mathrm{VOT}_i}{\sum_{i=0}^{n} \mathrm{weight}_i} \tag{8}$$

Mathematically, the elasticity of choice probability P of alternative A and attribute X of alternative B is given by:

$$E_{\mathrm{X_B}}^{P_i(\mathrm{A})} = \frac{\partial P_n(\mathrm{A})}{\partial \mathrm{X_B}} \frac{\mathrm{X_B}}{P_i(\mathrm{A})} \tag{9}$$

When aggregated over the population (similarly to Equation 8), this can be interpreted as the increase in choice probability for a unit increase in attribute X. To incorporate a habit analysis into this step, we segment and aggregate the elasticities by mode choice typology, to avoid endogeneity problems in the model specification arising from the inclusion of the habit strength variable which is highly correlated with the mode chosen. These elasticity values enable us to examine how changing one attribute could theoretically affect the modal split.

3 Results and Analysis

3.1 Overall mobility patterns

Below, we present a map of all trips observed in the sample (over the course of the study period), aggregated by postal code.

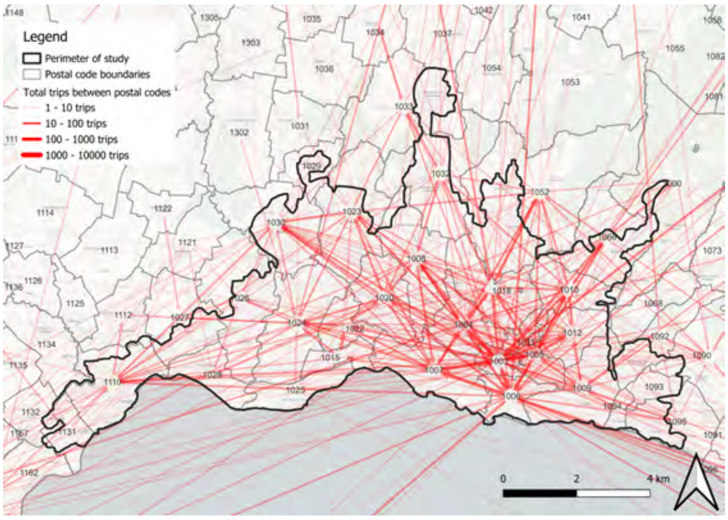
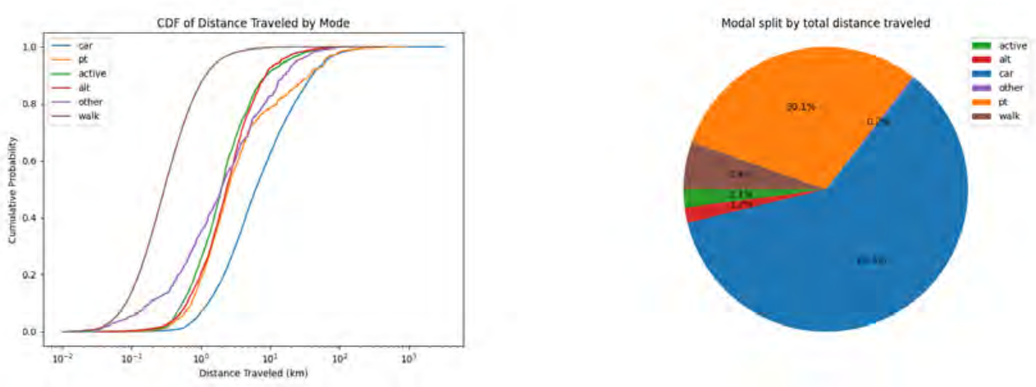


Figure 3-1: Origin/destination flows (by postal code) for the study perimeter.

As might be expected, we observe strong mobility into and out of the city center, with most mobility in the center occurring between adjacent zones. Few trips are generated between the eastern corner of the agglomeration and the center, with most longer-distance trips coming instead between West Lausanne and the center.

For additional context we provide below in Figure 3-2 the CDF of distance traveled by mode in the sample, as well as the modal split by distance traveled in the sample. For clarity, we exclude "Airplane" modes from the graphs below.



(a) Cumulative distribution of distance traveled by mode.

(b) Modal split in the sample by distance traveled.

Figure 3-2: Overall modal split tendencies, CDF (left) and overall (right).

The automobile is clearly the mode used to travel the farthest, with half of trips made by car being longer than 10km, and is subsequently the mode used for the vast majority (60%) of kilometers traveled in the sample. Behind the automobile is public transport, with 30% of the kilometers traveled. 80% of trips made by public transport are shorter than 4km (a similar rate as active, alternative, and other modes), and 10% are greater than 15km (a similar rate as automobile); only 10% of trips made by public transport are between 4km and 15km. Walking is responsible for 5.4% of kilometers traveled, but 90% of trips are less than 1km. Active and alternative modes are responsible for approximately 2% each of total distance traveled, but trips are generally used for longer distances than walking but shorter distances than by car. Finally, “other” modes (boat and other) are responsible for less than 1% of distance traveled.

3.2 Habit calculation

Here we present the results of the habit calculations. First, global results about the computed habits and a validation against the SRHI, and secondly presenting the results spatially with a comparison of the PT service quality.

3.2.1 Computed habits

In this section we first examine the overall results and compare with the SRHI, before discussing the clustering and segmentation results.

Overall results of habit strength by mode are presented in Figure 3-3.

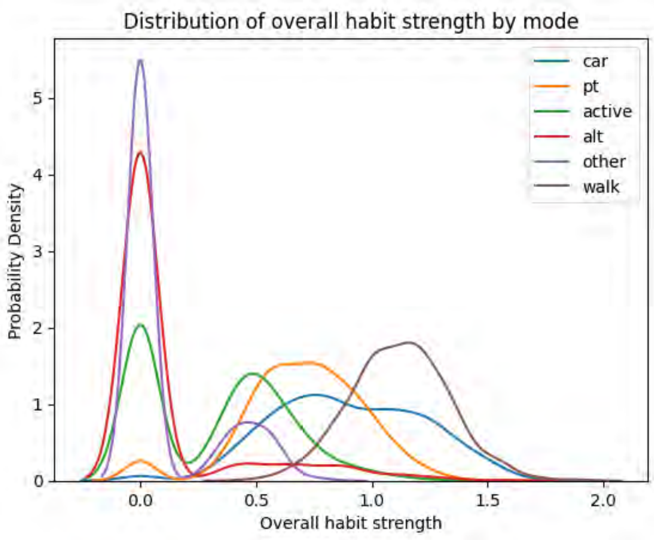
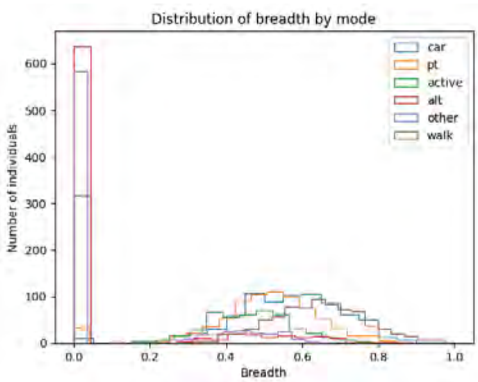


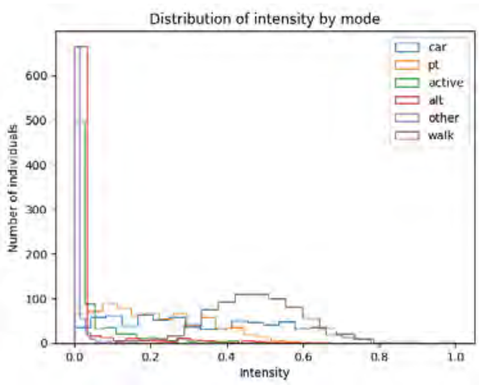
Figure 3-3: Probability density distribution of overall habit strength in the sample.

Most modes have a distribution with two peaks: one indicates those with zero habit (i.e. do not use the mode at all), and the other indicates those who use the mode at least once. Walking habit proved to be the highest throughout the entire population, with a mean population habit strength slightly above 1.0. Car habits are more variable within the population, and follow a wider distribution. The distribution of PT habits have a lower mean than car habits, but a narrower distribution, and a slightly higher proportion of individuals with zero habit. Active and alternative modes have approximately the same mean habit among those who use the mode, but slightly more individuals have an active mode habit. Alternatives have a very wide distribution, indicating a broad spectrum of habit strengths in the population. "Other" modes contain a high proportion of individuals with zero habit.

We continue with a presentation of the distribution of breadth and intensity values by mode in Figure 3-4:



(a) Distribution of habit breadth

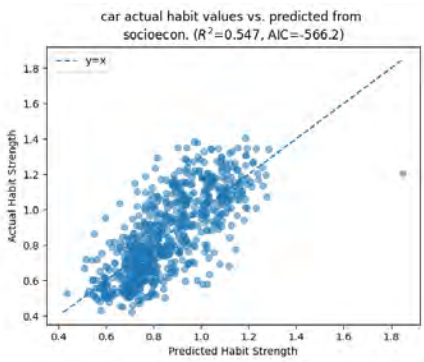


(b) Distribution of habit intensity

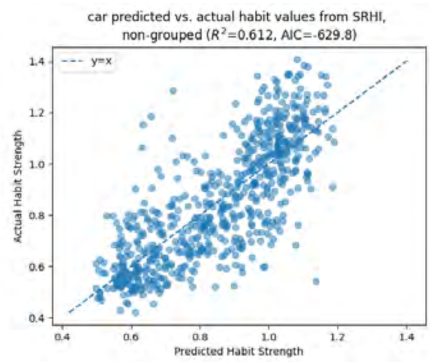
Figure 3-4: Overall distribution of habit breadth (left) and intensity (right) by mode.

Walking remains the mode with the highest breadth and intensity. Car and PT have similar breadth distributions–with car having a slightly higher mean–but the car tends to have a higher intensity than PT in the population. Active, alternative, and other modes have similar breadth distributions, but active modes have a greater portion of individuals with a habit. Alternative and other modes tend to have a higher intensity than active modes.

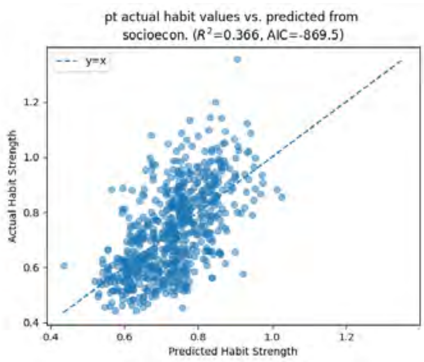
Comparison with SRHI To validate the results of the computed habits, we compare them with questions from the questionnaire similar to those found in the SRHI, as well as against a null model predicted from 25 purely socioeconomic variables such as age, gender, income, public transport passes, etc. The results are shown in Figure 3-5 below.



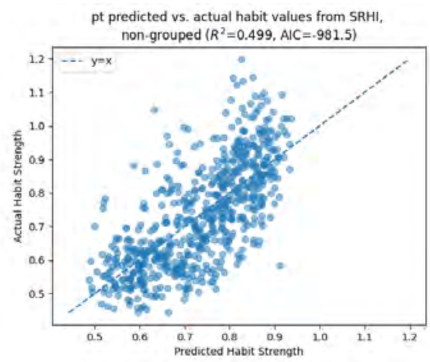
(a) Car habit predicted from socioeconomic



(b) Car habit predicted from pseudo-SRHI



(c) PT habit predicted from socioeconomic



(d) PT habit predicted from pseudo-SRHI

Figure 3-5: Regression of car and PT habits, predicted from socioeconomic data (left) and pseudo-SRHI (right).

The linear correlation coefficients and Akaike Information Criterion (AIC) for habit values predicted from SRHI (a combination of 4 aspects across 9 questions) demonstrated a significantly better fit ($p < 0.001$) than when predicting from 25 socio-economic factors.

3.2.2 Clustering

Hierarchical clustering was performed based on breadth and intensity values, using Ward linkage which builds groups based on the minimization of variance. A distance threshold for group segmentation was determined through the minimization of distance between observations within the groups in order to form a maximum of $g = 10$ groups; the resulting optimization yielded $g = 4$ groups. The dendrogram below shows the results of the hierarchical clustering at a calculated distance threshold of 6.0.

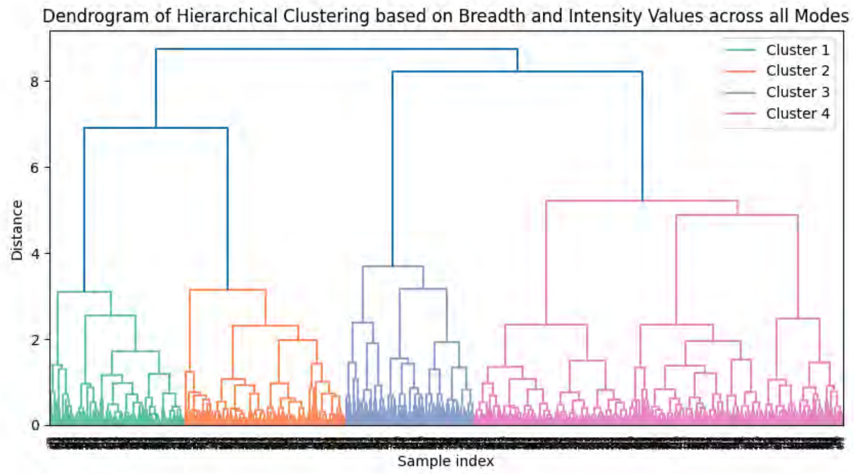


Figure 3-6: Hierarchical clustering dendrogram.

Cluster 1 and Cluster 2 are the first fully-formed clusters. They are more or less symmetrical (with the distance to the common ancestor being approximately equal for both clusters), and are clearly distinct from the other clusters. Clusters 3 and 4 are less symmetrical, with Cluster 4 containing three sub-clusters and having a slightly lower distance to the common ancestor than Cluster 3. It is also worth noting the relative proportions of each cluster in the sample, Cluster 4 being the most predominant with almost 50% of the sample:

Cluster	Individuals (Proportion)
1	131 (17.1%)
2	156 (20.3%)
3	123 (16.0%)
4	357 (46.5%)

Table 3-1: Relative proportions of each cluster in the sample.

Habit profiles created from Hierarchical Clustering also show distinct mode use patterns. To better visualize the results, we present in Figure 3-7 below the median overall habit strengths for each mode, by habit typology (clusters):

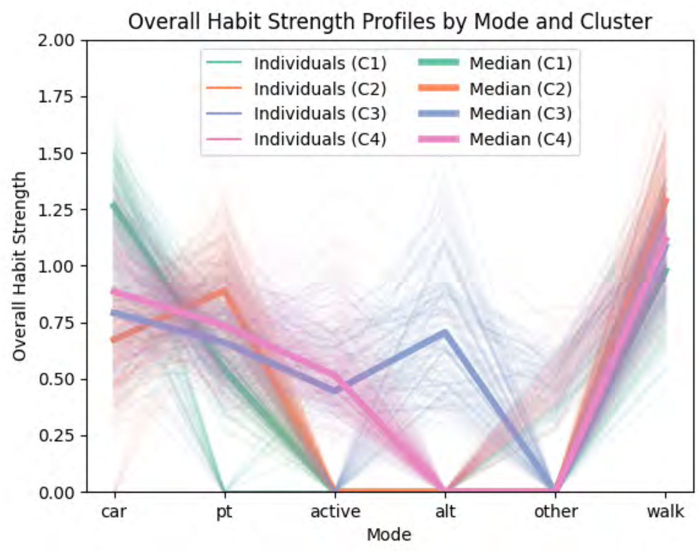
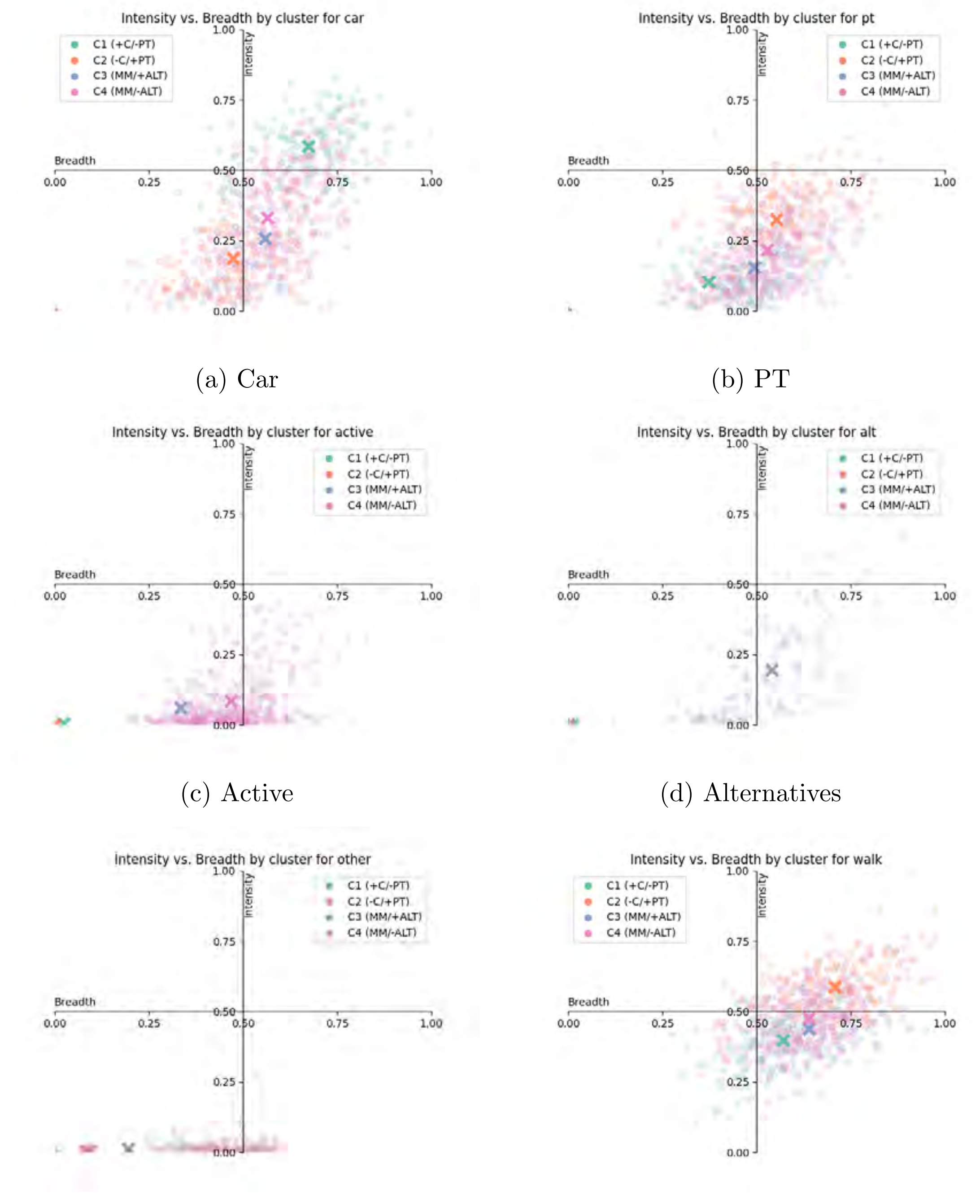


Figure 3-7: Overall habit strength profiles.

The clusters fall first into two broad categories based on the habit strength for active modes, with C1 & C2 tending to have low active mode habit strength, and C3 & C4 having a somewhat higher active habit strength. Within these two broad groups, the further distinctions are also quite clear. C1 and C2 are distinguished by their respective habit strengths for car and PT modes, with C1 having a strong car habit and a weak PT habit, and C2 having effectively the inverse, with strong PT habits and weaker car habits. Between C3 and C4, the key difference is in the active mode habit strength, with C3 having higher habit strength for alternatives (Motorbike and E-Bike) and C4 having generally no habit in the alternative modes. C3 and C4 can also be described as being multimodally-habituated, with neither tending to have very strong (> 1.0) habits in any one mode. Finally, all clusters tend to have high habit strength in walking.

Intensity vs. breadth plots (similar to Figure 1-6) for each mode by habit typology are presented below in Figure 3-8.



(a) Car

(b) PT

(c) Active

(d) Alternatives

(e) Other

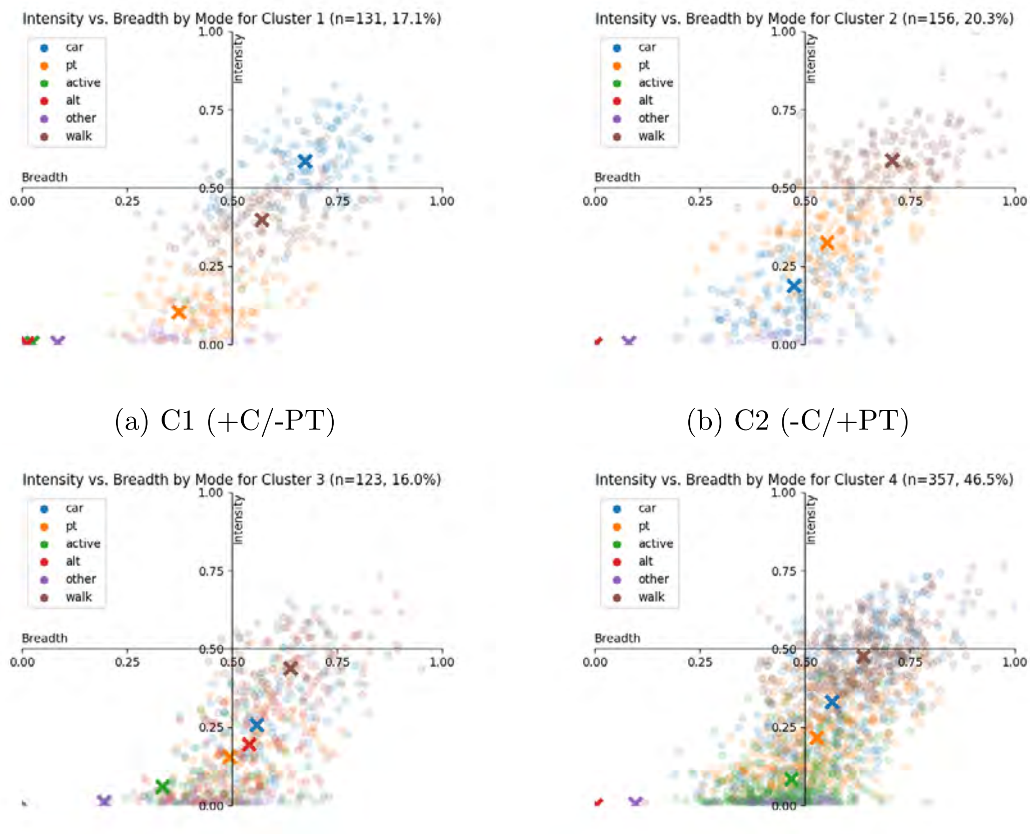
(f) Walk

Figure 3-8: Habit intensity vs. breadth plots per mode.

Car users in C1 have high intensity and high breadth car use habits, while car users in C3 and C4 are roughly equivalent in terms of moderate breadth, with C4 users having slightly more intense car habits; however, both intensities are nevertheless moderate. C2 users have even lower breadth and intensity values, with a mean breadth about twice that of the intensity. PT users in C2 have habits with moderate breadth and relatively low intensity, while for PT users in C1, mean habit breadth is somewhat low and intensity is quite low. Again, C3 and C4 users find themselves with both mean breadth and mean intensity values between C1 and C2, but with breadth values slightly closer to that of C2. Active mode users in C3 tend to have low breadth and low intensity, while those in

C4 have a higher breadth value. Active mode users in both groups, however, have a wide distribution of intensity values, with some individuals having upwards of 0.50. There are almost no active mode-habituated users in C1 or C2. Similarly, there are very few alternative mode users in C1 or C2, nor in C4. C3, however, boasts a moderate mean breadth, and a relatively low mean intensity in alternative modes. "Other" modes have few individuals with any habit, however, C3 individuals have a mean breadth value close to 0.20, with the other clusters having mean breadth values hovering close to 0.10; almost all individuals using "other" modes have very low intensity values. Finally, in walking, C2 individuals have the highest breadth and intensity values, both being close to 0.7. On the other end are car users, who have an intensity value below 0.5 and a breadth value close to 0.6. Mean breadth values of C3 and C4 are once more in-between C1 and C2, as are mean intensity values which are slightly lower and closer to that of C1. Users in C4 have slightly higher intensity values for the walking mode than those in C3.

It is also interesting to compare modal habits within clusters, so breadth-intensity plots by habit typology are also presented below in Figure 3-9:



(a) C1 (+C/-PT)

(b) C2 (-C/+PT)

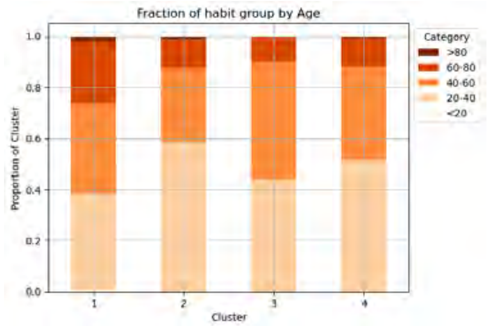
(c) C3 (MM/+ALT)

(d) C4 (MM/-ALT)

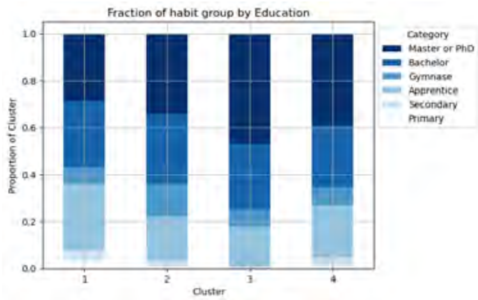
Figure 3-9: Habit intensity vs. breadth plots per habit cluster.

C1 is characterized by high car breadth and intensity, and low PT breadth and intensity, with walking somewhere in-between. C2 is characterized somewhat inversely, with moderate PT breadth and intensity, and relatively lower car breadth and intensity; walking is the mode with the highest breadth and intensity. C3 demonstrates no significant preference for one motorized mode over others, however, car and alternative modes have approximately equal mean breadth, and the mean intensity of the car habit is slightly higher than that of the PT mode; walking represents the mode with the highest mean breadth, and a moderate mean intensity. C4 individuals have a relatively high active mode habit breadth, albeit lower than PT and subsequently car; the car is the motorized mode with the highest intensity, and walking is used at an overall modest intensity with relatively high breadth.

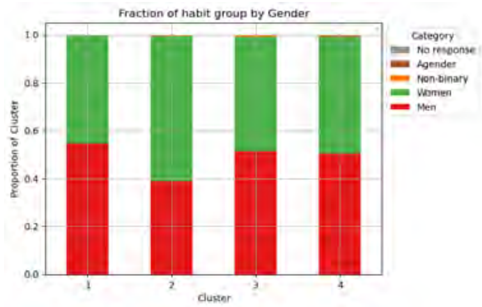
Socioeconomic characteristics of each cluster were also examined, and are presented in Figure 3-10 below:



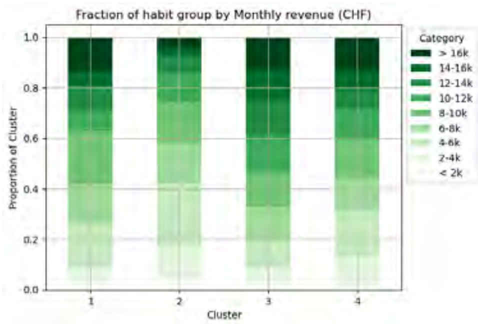
(a) Age



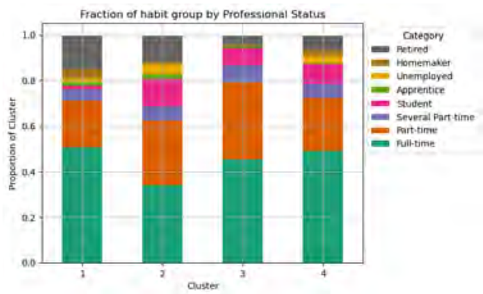
(b) Education



(c) Gender



(d) Income



(e) Professional status

Figure 3-10: Socioeconomic characteristics of clusters.

Regarding age, C1 comprises the largest share of people over the age of 60 (25%), while C3 comprises the smallest (10%). C2 is the youngest group, with 60% of its members younger than 40.

By education, C3 is the highest-educated group, with 75% of the group having earned at least a Bachelor's degree. C1 has the highest proportion having attained primary or secondary school (4.2% each), as well as the highest proportion of those having completed an apprenticeship (30% vs. 20% in other groups). C2 is the group with the highest proportion having completed gymnase (11% vs. 6% in other groups) as well as having completed a Bachelor's degree (29%).

In terms of gender, C1 comprises the highest proportion of men (54%), while C2 has the greatest share of women (57%). Examining this breakdown by age is even more telling: 63% of individuals in C2 younger than 40 are women, representing the strongest age split of any other cluster. While in C1, C2, and C3 the gender split in the 60-80 group favors men (approx. 57% men/43% women), in C4 the trend is reversed, with women being the majority at 54% against 46% men. Similarly, young men dominate the under-40 age group at 60% in C1, where women are the majority in the other clusters. Finally, in the 40-60 age range, the majority are women in C1 and C2 (57% and 64% respectively), while men are the majority in C3 and C4 (56% and 57% respectively).

C3 is the wealthiest group, with the 43% of individuals earning more than 10,000 CHF/month. Conversely, C2 is the least wealthy group, having both the highest proportion of individuals earning less than 10,000 CHF/month (80%), as well as the lowest proportion of individuals earning more than 10,000 CHF/month (20%).

Finally, by employment, C1 contains the greatest proportion of homemakers (3.8%) and retirees (25%) relative to the other groups. C1 also has the greatest share of young people less than 40 holding full-time jobs (64%) as well as the lowest percent of young people holding one (22%) or more (4.3%) part-time jobs. C2 is the group with the highest proportion of students (8%), with 20% of individuals younger than 40 in this group engaged as a student. Contrarily, C1 is the group with the lowest proportion of students (< 1%), as well as with the lowest proportion of individuals with one or more part-time jobs (20% and 4.6%, respectively). C3 has the lowest portion of unemployed (0%), homemakers (0.5%) and retirees (7.6%), and the highest percentage of individuals with one (32%) or more (8%) part-time jobs. Finally, C4 is the group with the overall highest proportion of full-time workers, with 49% of the group holding full-time employment.

Modal attitudes are also examined, as the Panel data included questions about perceived mode suitability for different trip types, whose responses we segment by habit typology. The results of this segmentation are shown in Figure 3-11 below; bars for a given plot indicate the average tendency of individuals in a given typology to perceive the given mode as appropriate for the corresponding trip category.

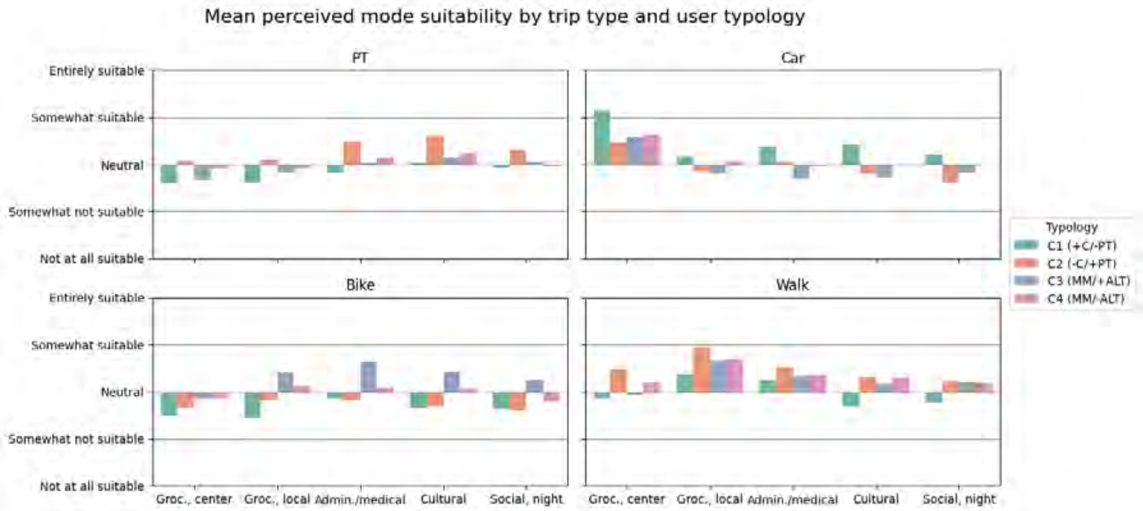


Figure 3-11: Perceived mode suitability by mode, trip type, and habit typology.

Figure 3-11 above shows that C2 is the group that perceives PT the most suitably across all categories, among all groups. Similarly, C1 is the group that perceives the automobile as the most suitable across all trip categories, among all groups. C3 is the group which perceives the bicycle as the most suitable for the most trips, among all groups. Finally C4 lies in-between the others, with a generally higher perception of PT suitability than C2, a more moderate perception of car suitability than C1, and a more moderate perception of bicycle suitability than C3. The only activity for this group in which the car is perceived as more than neutrally is for grocery shopping in a commercial center. This is a similar attitude as the C2 group, but the perception of PT in this group is less suitable than C2.

We also examine the degree of comfort using public transport segmented by modes, below in Figure 3-12:

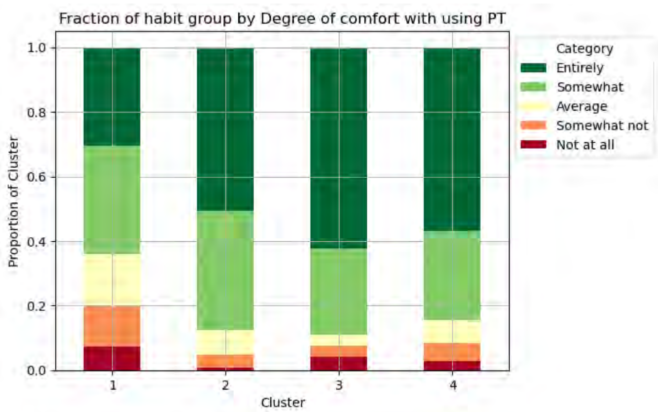


Figure 3-12: Cluster composition by comfort using PT.

C2 comprises the least proportion of individuals who are at all uncomfortable with public transport (2%) while C1 comprises the highest proportion of people who are at all uncomfortable with public transport (9%). Of the individuals in C1 who are not at all comfortable with public transport, two-thirds are men, of whom 50% are younger than 40. The similar pattern holds for group C3, where 60% of those not at all comfortable taking PT are men, of whom 100% are under 40. However, the majority of those in C4 who are not at all comfortable taking PT are women, at 56%, of whom 80% are under 40.

Finally, we examine below in Figure 3-13 the modal split in Lausanne (by distance traveled), segmented further by habit typology:

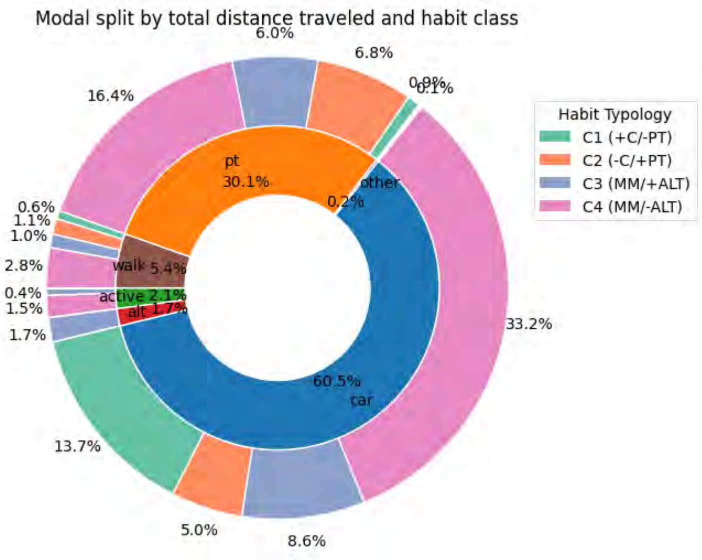
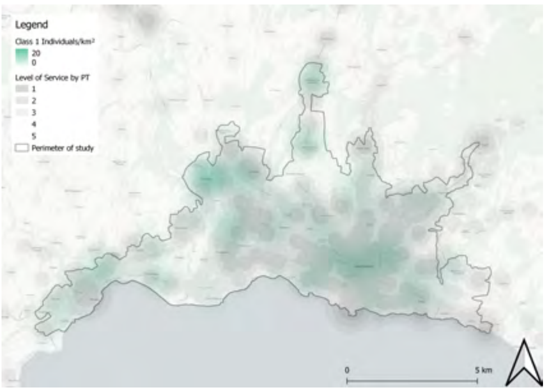


Figure 3-13: Modal split by typologies.

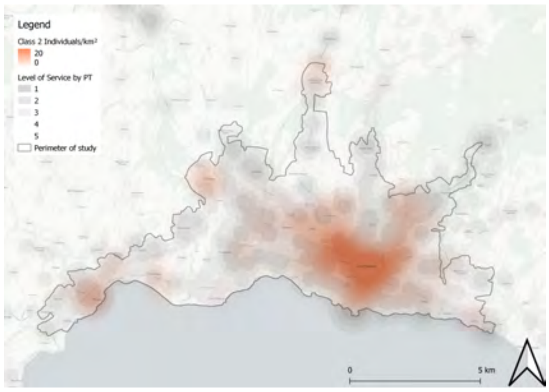
The majority of distance traveled in most modes was done by individuals in C4, generally accounting for at least 50% of distance traveled. The exception is alternatives, which are exclusively dominated by C3 individuals. The portion of car-kilometers traveled by individuals in C1 is disproportionate relative to their share of the population: these individuals account for 23% of the distance traveled by car, but only 17% of the population. The same is true of C4 individuals, who account for 46% of the population, but 55% of the distance traveled by car. In fact, in general, these C4 individuals tend overall to accrue more kilometers across all modes, accounting for 54% of all kilometers traveled (even though they comprise 46% of the population). C1 individuals also account for a disproportionately small share of PT and walking kilometers, just less than 1% and 11%, respectively. 100% of alternative mode kilometers are traveled by C3 individuals.

3.3 Spatial analysis and service gaps

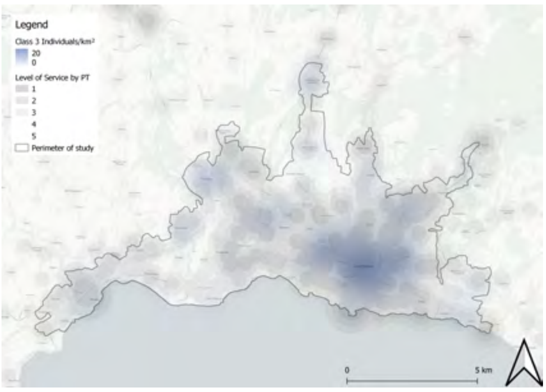
Here we present an analysis of the spatial distribution of the different habit clusters, and correlate these distributions with the Vaudois index of PT service quality. Figure 3-14 below presents kernel density estimations of the inferred home locations for each habit cluster.



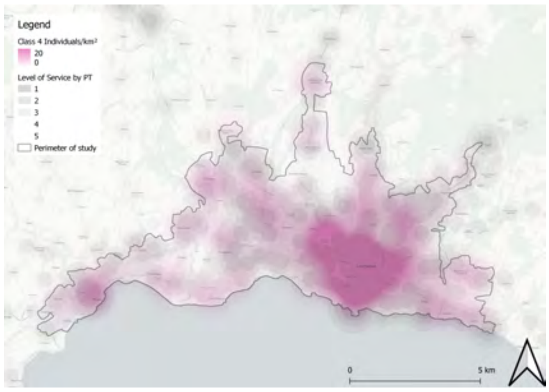
(a) Cluster 1 (+C/-PT)



(b) Cluster 2 (-C/+PT)



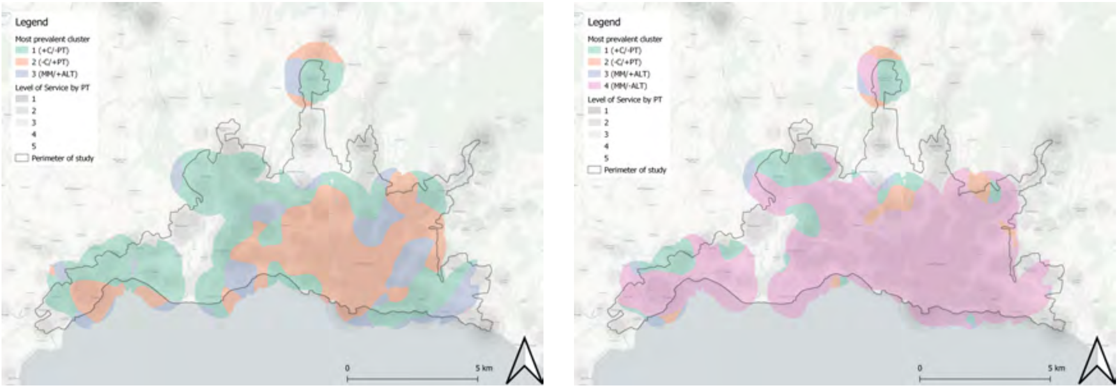
(c) Cluster 3 (MM/+ALT)



(d) Cluster 4 (MM/-ALT)

Figure 3-14: Kernel density estimations (indiv./km^2) of inferred home locations for each habit cluster, superimposed on the index of public transport service quality.

From the maps above, we observe very strong spatial distribution patterns: C4 (MM/-ALT) is strongly distributed throughout the entire agglomeration, with the strongest mass concentrated in Lausanne Proper. We observe similar, but weaker, concentrations of C2 (-C/+PT) and C3 (MM/+ALT) in the city center, however, C3 is more homogeneously distributed throughout the entire agglomeration. Finally, C1 (+C/-PT) is the least concentrated, but distinct pockets are distinguishable in the peripheral centers. To make these spatial patterns more clear, we present also in Figure 3-15 a map of the cluster with the highest concentration in each area. For clarity, in Figure 3-15a we exclude C4, which tends to dominate everywhere.



(a) Only clusters 1-3 (b) All clusters

Figure 3-15: Spatial predominance of mode habit clusters.

Figure 3-15a above shows the spatial predominance of clusters 1-3. PT-habituated individuals (C2) are the most common group in urban centers, while car-habituated individuals (C1) are concentrated in the periphery. Alternative-preferential multimodal individuals (C3) are most prevalent in-between pockets of these two groups, and where PT service quality is lower.

In Figure 3-16 below, we present the distribution of public transportation level of service at the inferred home locations according to habit cluster:

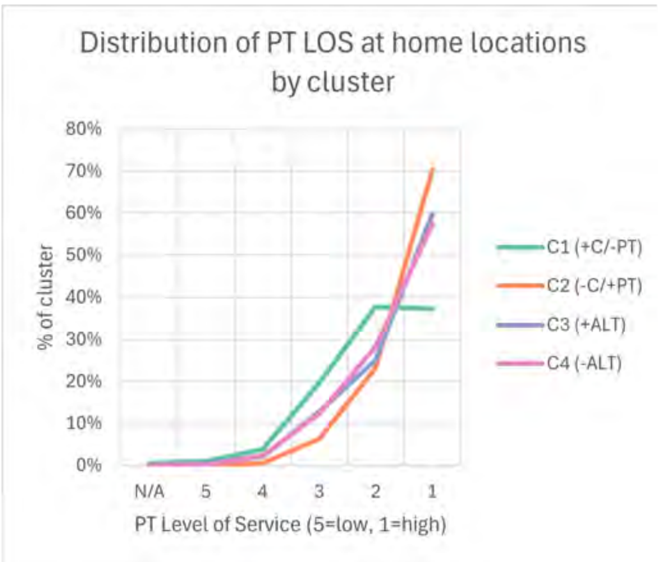


Figure 3-16: Distribution of PT service quality at home locations, by cluster.

Individuals in the PT-preferential group tend to live in areas of high PT service quality, while individuals in the car-preferential group tend to live in areas of low PT service quality. Multimodal individuals (clusters 3 and 4) are relatively indistinguishable from one another, but are in-between the other clusters in terms of PT service quality.

3.4 Mode choice model

3.4.1 Calibration of NPVM to Lausanne

The calibration of the model to the dataset was performed using the open-source Biogeme package in Python [38]. Full parameter estimates can be found in Table A-5. The final model had 31 parameters and a final log-likelihood of -3521.58.

3.4.2 Value of time

The final aggregate population VOT was estimated to be XXX CHF/hr for PT and XXX CHF/hr for car. Figure 3-17 below shows the distribution of individual VOT in the subsample for PT and car:

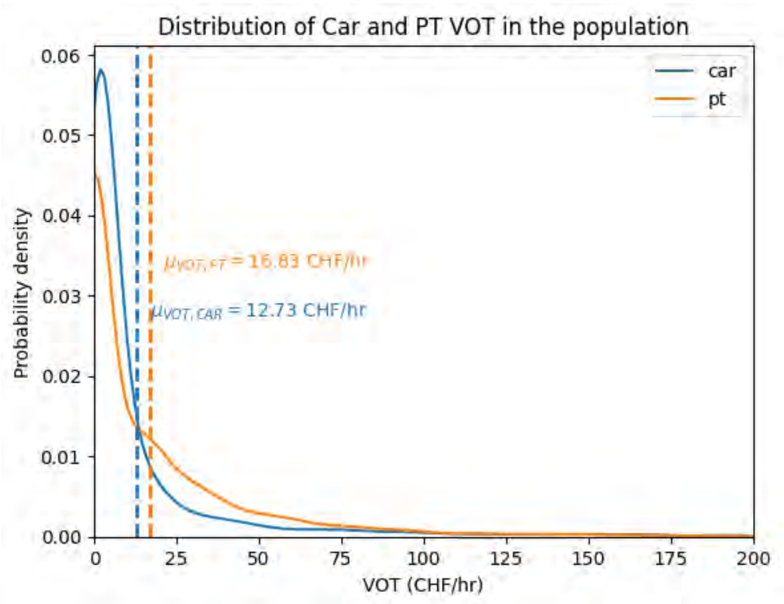


Figure 3-17: Distribution of VOT in the subsample.

The VOT for car was found to be slightly lower than that of PT, at 12.73 CHF/hr for car vs 16.83 CHF/hr for PT.

3.4.3 Mode choice probability elasticities

Mode choice probability elasticities were calculated for all mode attributes included in the model specification. Figure 3-18 below presents the elasticity results graphically, segmented by habit typology:

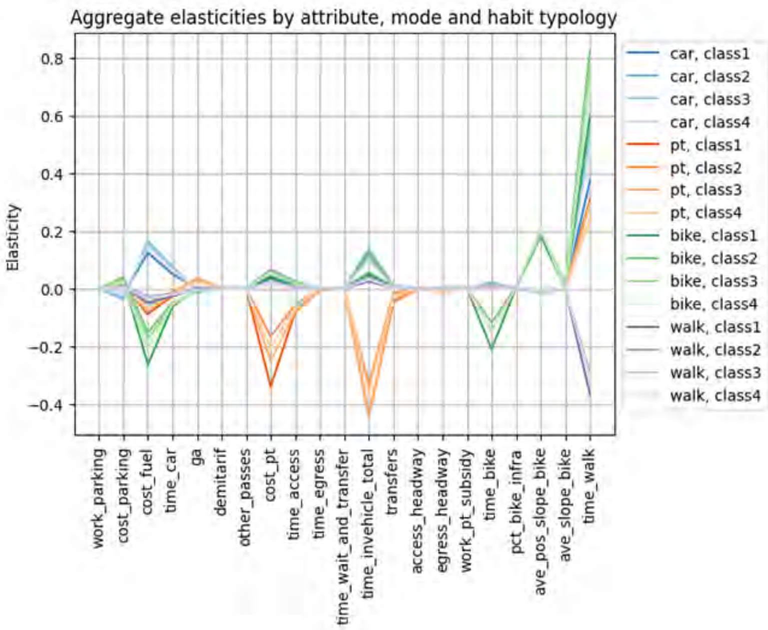


Figure 3-18: Distribution of mode choice probability elasticities, segmented by habit typology.

Positive values in the graph above suggest a positive relationship between the associated variable and mode choice probability. The model predicted a positive elasticity for driving time and fuel cost, but negative elasticities for PT cost and travel time components. The class that was least sensitive to changes in fuel cost and travel time is C1, which also was the group most sensitive to PT cost. Car choice was also little affected by changes in the active mode attributes, across the board. The bike mode was generally the most sensitive to perturbations, with C1 individuals being the most sensitive subgroup. Little distinction is obvious between classes regarding bike choice probability for average positive slope. Choice probabilities for bike are most sensitive to walking time, with C3 and C4 individuals experiencing strong increases in bike choice with increasing walking time. Public transport choice was least sensitive to walking time, with C3 and C4 individuals being the least sensitive subgroup.

4 Discussion

4.1 Calculated habits

The results of this newly-proposed method to quantify habits has proven to correspond well across several indicators. Firstly, habits calculated from this method were significantly better predicted from the pseudo-SRHI than from socioeconomic variables. Thus we can reject our H_0, that our pseudo-SRHI is no better at predicting calculated habit strength than purely socioeconomic variables, which validates the proposed method of habit strength against the existing SRHI. Granted, our new metric relies on the implicit capture of the aspects captured explicitly in the SRHI, but we do nevertheless observe that the SRHI is better able to predict habits than socioeconomic variables, and with less variables. It should be noted that calibration is theoretically possible using maximum likelihood estimation instead of numerical simulation, but this implementation was not explored here due to time constraints.

Secondly, the calculated habits also correspond with self-reported modal attitudes. We observe that individuals who are habituated with public transport (C2) are more comfortable with it as a transportation mode, as well as the group that perceives it as the most suitable for all situations examined. The only group with a car habit stronger than walking habit (C1) is also the group that reports perceiving cars as the most suitable and walking as the least suitable mode for almost all situations examined. In addition, calculated breadth also appears to roughly correspond with perceived mode suitability; the more similarly a mode's suitability was scored across trip types, the wider the mode's mean breadth value.

Another interesting result of the intensity/breadth analysis is that very few individuals possessed a high intensity, low breadth habit. This could be a consequence of unbalanced metrics (intensity tends to skew lower as a result of the quadratic mean). Further study could be taken to more rigorously define these metrics in such a way that ensures their balance. This suggests that modal tendencies tend to bleed across trip categories; if an individual uses a mode in a given context, they are likely to use it for several, as the efficacy of this mode in this context is reinforced through repetition. Certainly, habits are context-defined, but these results suggest that, with repetition and reinforcement, habits may even transcend contexts.

The overall modal split of the sample corresponds to the most recent estimates from the Ville de Lausanne: 60% by car, 30% by PT, 2% each by active modes and alternatives, and 5% by foot (against 45-59% by car, 22-38% by PT, 3-6% by bike, 11-35% by foot, respectively, in 2022 [2]), within which, there were some further tendencies relating to habits. We observe that those with strong car habits are responsible for a disproportionately greater share of kilometers traveled by car and a disproportionately small share of kilometers traveled by public transport. Similarly, those with strong PT habits are responsible for a disproportionately greater proportion of kilometers traveled by PT as well as a disproportionately smaller proportion of kilometers traveled by car. Individuals with multimodal habits also tend to account for a significant portion of kilometers traveled across all mode categories. These individuals may prove to be crucial in driving the modal shift, as they accumulate more kilometers of travel than the average individual, but are not necessarily faithful to one mode in particular. This is in-line with the findings of [4] presented in Section 1.2.1, which showed that 80% of car traffic in Lausanne was generated by daily automobiles users who only count for 41% of the population despite a mere 8% of the population being "car-exclusive".

4.2 Typologies

Car-habituated individuals (C1) dominate all other clusters in Bussigny and north of Ecublens, Echichens, Cheseaux-sur-Lausanne, and Jouxtens between the A9 highway and Prilly; in these pockets on the edges of the agglomeration, autoroute accessibility is high, while LOS higher than 2 is rare. Further, C1 dominates all clusters other than C4 in the smaller, peripheral urban centers of Prilly, Preverenges, and Renens. These individuals use the car frequently and in a broad range of situations (high breadth, high intensity), complementing their mobility with walking, and using public transport very rarely, only in specific situations (low breadth, low intensity). These individuals thus use the automobile almost per default, choosing it over PT in almost every situation. There are nevertheless some situations for which they would use PT, but not unless it was necessary. This likely stems from these individuals' relationship with public transport, with which this group feels the most discomfort relative to the other groups. These inferences from breadth and intensity are confirmed by Figure 3-11, which show that for shorter, smaller local trips with less need to transport large items, the car is perceived as less suitable, but walking is perceived as more suitable than PT in these situations. Individuals in this group may not want to make the journey back home in the periphery via public transport after a night out or a large shopping trip into the city center.

Public transport-habituated individuals (C2), on the other hand, rarely dominate C4. Instead, they dominate all other clusters, directly in the urban centers, as well as around the universities EPFL and UNIL in Ecublens. Interestingly, this cluster's highest breadth and intensity are for walking, which is complemented by a PT habit of moderately high intensity and moderately low breadth. This means that walking is almost always taken at least as often as public transport in at least as many situations, which is explained by the necessity of walking either to access, egress, or transfer with PT. The car in this group is used in even less situations, yet to a greater extent than C1 uses PT (the exceptional case is grocery shopping in a shopping center, for which the car is perceived as well-adapted, likely due to the need to transport items a long way from the shopping center in the city back home to the periphery). These individuals thus perceive the car as a useful and viable alternative to PT in several situations, but PT is given the priority. For students, with low incomes and flexible schedules, PT makes much more sense than car ownership, being more easily able to afford a ticket than a few hours in a parking garage. This modal script also makes sense for the large share of non-full-time workers in this group, whose time budgets tend to be greater as a result of less strict and/or less regular schedules. Further, students are able to benefit from a reduction in PT costs through a subsidy from the Ville de Lausanne. Finally, other costs associated with car use may also dissuade these individuals from frequent car use due to these individuals' tendentially lower income.

Multimodal-habituated individuals with a preference for alternatives (C3) appear in clusters in-between pockets of C1 and C2, generally in pockets in-between or on the border of zones with relatively high PT LOS. Examining the breadth-intensity graphs in Figure 3-9c, we observe that walking has the highest mean breadth and intensity, while the car, alternatives, and PT have approximately similar breadth and intensity values, with car tending to have slightly higher breadth values and higher intensity than the alternatives, which have a higher intensity and breadth the PT. This confirms that alternatives are in fact used to fill a perceived gap between automobile and TP, given that the mean breadth (and intensity) of the car habit is much less than that of C1, the mean breadth of the PT habit is greater than that of C1, and the strength of the mean breadth of alternatives is much higher. These indviduals also complement their activities with active modes, but at a much lower intensity, and lower breadth than the other modes, suggesting it to be more of an exception than a rule.

Multimodal-habituated individuals with a low preference for alternatives (C4) are the most dominant cluster almost everywhere in the sample perimeter. Walking is used to fill many everyday mobility needs at a moderate intensity, followed by car use and then public transport. Complementing this core mobility are active modes, used infrequently in a moderate number of daily situations. This represents the highest breadth of active mode use in in any class. Public transport habit strength is higher than with C3, as car habit, while the use of alternatives is entirely absent. This suggests that in situations where C3 individuals would turn to alternatives, C4 individuals are more likely to remain faithful to the car and then PT, opting instead for active modes when neither of these are feasible. However, by stated suitability, the car is only perceived as more than neutrally suitable for one situation, and PT perception is close to neutral. It is possible that this represents a lack of information or education regarding the efficacy of PT, as these individuals are only slightly more uncomfortable with PT than those in C2. Surprisingly, these individuals do not appear to perceive the bike as suitable for many situations, yet this group is responsible for the greatest share of active mode kilometers, as well as the highest habit strength for this mode. Much of this could also be explained by the fact that these individuals tend to be full-time workers, whose schedules are generally fixed and require more effort to adapt to PT timetables and potential service perturbations.

4.3 Behavior/planning gaps

We also observed strong patterns in the spatial predominance of clusters. Patterns similar to [5] were found, where the concentration of individuals with a preference for automobiles (C1) tends to increase with distance from the city center, those with preference for alternatives (C3) are most concentrated in the city center, and those with multimodal tendencies (C4) are distributed overall. C3 individuals, exhibiting an alternative mode habit and typically situated in areas between zones of high public transport service, appear to perceive alternatives as a viable mode for filling a perceived service/accessibility gap between car and public transport, while having a moderate habit strength for all three modes. In these areas, where individuals are close to agnostic between car and PT modes, targeted improvements to public transport connectivity or accessibility in these areas might be just enough to tip the scales away from the car.

Examining Figure 3-16, we notice an abrupt plateau in the proportion of C1 (+C/-PT) individuals with increasing LOS around LOS 2, suggesting that they are not as drawn to living near high public transport service as the other groups, and thus increasing PT LOS may not be as convincing for these groups. Interestingly, despite C4 having a stronger PT habit, the distribution of C3 and C4 individuals across LOS is approximately the same. Conversely, PT LOS has a strong relationship with home location for individuals with a strong PT habit.

This confirms the work of several transportation researchers in recent years seeking to go beyond mode attributes and model more exactly this relationship between land use and mode choice. Notably Schmid et al. (2023) found strong relationships between home location, attitudes, and mode choice, with the concession that the model could not account for behavioral/attitudinal shifts [39]. Importantly, the same paper found that a simpler model, only considering modal attributes and excluding attitudes and residential location, drastically overestimated modal shift dynamics compared to a three-dimensional one. This is likely in part due to the aforementioned issue of such one-dimensional models not accounting for attitude shifts, as well as a lack of longer-term influences on choice (i.e. residential location, habits, etc.). With this important context in mind, we present finally the results of the logit mode choice model calibrated to Lausanne.

4.4 Mode choice model

It should first be noted that the model results obtained are far from robust, as a result of two main limiting factors: the alternative generation procedure and the multicollinearity of travel time and travel cost. The former was limited by time and computational restraints, which prevented all alternatives from being generated from the same engine.

The second limiting factor stems from the fact that travel cost and travel time are often highly correlated, which results in numerical instability issues and nonsensical values, such as the positive coefficient of both travel time and fuel for the car mode. This, however, remains a problem with these types of mode choice models to this day [40].

Nevertheless, even with this faulty model, we found that mode choice elasticities did vary by habit class, generally in a way that corresponded to our expectations. Individuals with strong modal habits were generally less sensitive to changes in the attributes of that mode, and individuals with weaker habits were generally more sensitive to changes in attributes of that mode. On the contrary, sensitivity to fuel cost and time were inverse from expected values, as a result of the aforementioned multicollinearity issue.

This brings us to a discussion on the general utility of such models. The mode choice model is unable to explicitly capture any habitual behavior, and assumes that individual preferences and travel behavior remain static in response to an imposed policy change. Of course, this is not actually the case. However, what we may be able to take from such models is a general sentiment according to population segments. This can also be a gauge for how individuals with low modal habits might respond to such a change, with these individuals posing the least resistance. We might additionally imagine that the greater the predicted modal shift, the more complimentary effort required on the part of the policymaker to overcome the resistance posed by individuals with strong habits. Being able to examine the proportion of individuals expected to shift who have strong modal habits can give some indication as to which populations should receive additional sensibilization efforts.

These results are an interesting exercise in the discussion of the role of habits in planning tools. Fundamentally, such mode choice models are based on an assumption of perfect omniscience, rationality, and optimization of self-interest as well as that of “choice”. We have seen, however, that these assumptions do not hold in reality, and that the concept of “mode choice” in everyday life is more driven by habit than by rationality. This is not to say that these models are not useful; indeed they provide insight into correlations between tendency to use a particular mode, its attributes, and an individual’s socioeconomic characteristics. However, such correlations unable to provide explanatory power. This distinction is particularly important in the development of modal shift policies, which should be based on causal relationships to effectively incite individuals to change their behavior.

5 Conclusions

This paper takes another step forward in the field of operationalizing modal habits in mobility research. Substantial theoretical groundwork has been done to date, but due to the complexity of the problem space, there has been little attempt to operationalize these theoretical concepts. This research positions itself between planning tools and individual behavior in an attempt to bridge the theory/operational gap. Building on previous successes in the prediction of daily mobility [18], a Hidden Markov Model approach was selected for its flexibility and suitability to the problem. The resulting calculated habit strengths showed strong socioeconomic and spatial patterns, and proved in-line with aspects of the current leading metric of habit strength. Incorporating habits into analyses of public transport quality as well as model predictions from a classical model offer potentially more nuanced planning decisions.

To operationalize the work presented here, given that habits are as much socially constructed, efforts for driving a modal shift might be best spatially targeted at intersections of spatial dominance boundaries. That is, if individuals living near these spatial dominance boundaries begin to start gently adapting their habits in response to an injunction, these new behaviors may gradually manifest into full-fledged habits, which other individuals in the neighborhood may also remark on and adapt their behavior accordingly. One of the most important areas in this sense is the corridor between Renens and Prilly, which comprises a pocket of strongly-PT habituated individuals dominated to the north by car-habituated individuals and to the south by multimodal individuals with no alternative mode habit. These multimodal individuals, despite not having a significant car habit, are disproportionately responsible for kilometers traveled by car. Targeting policies in this region specifically to promote public transport over the car could have significant impact on preventing these openminded individuals from developing a stronger car habit, and persuade them to take other modes.

Finally, regarding overcoming modal habits in policymaking, Thomas Buhler makes the distinction between two main types of policies: implicit and explicit injunctions. The first type considers the individual as an instrumental actor in urban transports, meaning that the message *implied* by a given policy (i.e. reduction of parking places in the city center, or increased PT offer) are assumed to be conveyed and interpreted by the users, simply be means of the existence of the infrastructural changes. The second type considers the individual as an axiological actor, who acts according to values attributed to associated behaviors; policies in this category appeal to the logical chain of information, sensitization, and education. The form of this second type of policy reminds us of the chain of schema internalization mentioned earlier in Section 1.3.1: the internalization of a new habit can occur only in a state of *ethos*, or when an individual is receptive to new schemas. According to Buhler, once habits are formed, being receptive to new schemas is rare, and comes usually in the form of "decisive moments," when life circumstances suddenly change (a new job, moving, retirement etc.), because these are moments "where the individual is less engaged in a habituated process, and where decisions can be made based on associated values or objective qualities of certain modes," [7, p.109, own translation]. In this light, effective planning for the modal shift can be done to take advantage of these "decisive moments," and target individuals on the verge of a major change in life circumstances. This approach can also benefit from these habit typologies, as we have seen from the sample that younger individuals, students, and part-time workers tend to be more open to alternative modes, while older individuals, full-time workers, and retirees dominate the group with strong car habits. Thus captializing on naturally-occurring demographic shifts could also play an instrumental role in the adaptation of behavior.

Despite the positive outcomes, this paper is not without its limitations. For one, the metrics devised for habit breadth and intensity are unbalanced (marginal breadth is not strictly equivalent to marginal intensity) and are not necessarily orthogonal (resulting in potential information overlap). Future work could use more rigorous statistical methods to improve upon these metrics. However, this could also simply be a result of particularities in the dataset or location, an issue which could be further confirmed by attempting to generalize the approach to less detailed datasets. Results from the mode choice model should also be taken lightly; time and computational constraints limited the cohesiveness and robustness of the alternatives generated, and a multicollinearity problem between travel time and travel cost proved very difficult to resolve. Once resolved, further analysis could attempt to draw more robust relationships between habit strength and elasticities, rounding out this initial attempt. This further underscores the gap between existing planning tools and individual behavior, emphasizing a significant research gap in incorporating habits into more complex models of mode shift dynamics.

We recall once again the goal of this project, which was to establish a new metric to quantify modal habit strength, and use this metric to identify gaps between theory, practice, and actual human behavior. Although not perfect, this work aims to simply be an exploration of what is possible, serving as a first step at reconsidering the role of modal habits in overcoming resistance to the modal shift.

References

[1] European Parliament, "CO2 emissions from cars: facts and figures (infographics)," https://www.europarl.europa.eu/topics/en/article/20190313STO31218/co2-emissions-from-cars-facts-and-figures-infographics, accessed: 2024-06-04.

[2] Ville de Lausanne, "Observatoire de la mobilité 2023," Service de la mobilité et de l'aménagement des éspaces publics, Lausanne, Switz, Tech. Rep., Jun. 2023.

[3] Federal Office for Spatial Development (ARE), "Population's travel behaviour 2015," Federal Statistical Office (OFS), Neuchâtel, Switz, Rep. OFS 1697-1500, 2017.

[4] V. Kaufmann, J. González, E. Bernier *et al.*, "Analyse des logiques de choix modal auprès de la population active urbaine : Etude comparée du grand genève, du canton de vaud et des agglomérations de berne et de bienne," Laboratoire de Sociologie urbaine (LaSUR), Lausanne, Switz, Cahier du LaSUR 33E, 2020.

[5] ——, "Analyse des logiques de choix modal auprès de la population active vaudoise," Laboratoire de Sociologie urbaine (LaSUR), Lausanne, Switz, Cahier du LaSUR 33C, 2019.

[6] Federal Office for Spatial Development (ARE), "Comportement de la population en matière de mobilité: résultats de microcensement mobilité et transports 2021," Federal Statistical Office (FSO), Neuchâtel, Switz, Rep. OFS 841-2100, 2023.

[7] T. Buhler, *Déplacements urbains: sortir de l'orthodoxie : Plaidoyer pour une prise en compte des habitudes.* Lausanne, Switz: Presses polytechniques et universitaires romandes, 2015.

[8] C. LeMieux, "Problématiser," in *L'enquête sociologique.* Presses Universitaires de France, 2012, pp. 27–51.

[9] P. Rietveld, "Six reasons why supply-oriented indicators systematically overestimate service quality in public transport," *Transport Reviews*, vol. 25, no. 3, pp. 319–328, 2005.

[10] T. Thomas, L. Puello, and K. Geurs, "Intrapersonal mode choice variation: Evidence from a four-week smartphone-based travel survey in the netherlands," *Journal of Transport Geography*, vol. 76, pp. 287–300, 2019.

[11] L. Redman, M. Friman, T. Garling *et al.*, "Quality attributes of public transport that attract car users: A research review," *Transport Policy*, vol. 25, pp. 119–127, 2010.

[12] Aristotle, *Nicomachean Ethics.* Cambridge University Press, 2000.

[13] G. S. Becker and K. M. Murphy, "A theory of rational addiction," *Journal of Political Economy*, vol. 1988, no. 4, p. 26, 1988.

[14] B. Lahire, *L'homme pluriel : les ressorts de l'action.* Paris: Nathan, 1998.

[15] G. M. Hogdson, "Choice, habit and evolution," *Journal of Evolutionary Economics*, vol. 20, no. 1, pp. 1–18, 2010.

[16] R. Poldrack, J. Clark, , J. Pare-Blagoev, D. Shohamy, J. Creso-Moyano, C. Myers, and M. Gluck, "Interactive memory systems in the human brain," *Nature*, vol. 414, no. 29, pp. 546–550, 2001.

[17] B. Verplanken and S. Orbell, "Reflections on past behavior: a self-reported index of habit strength," *Journal of Applied Social Psychology*, vol. 33, no. 6, pp. 1313–1330, 2003.

[18] Q. Lv, Y. Qiao, N. Ansari, J. Liu, and J. Yang, "Big Data driven hidden Markov model based individual mobility prediction at points of interest," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 5204–5216, 2017.

[19] R. Su, E. McBride, and K. Goulias, "Unveiling daily activity pattern differences between telecommuters and commuters using human mobility motifs and sequence analysis," *Transportation Research Part A*, vol. 147, pp. 106–132, 2021.

[20] C. Schneider, V. Belik, T. Couronné *et al.*, "Unravelling daily human mobility motifs," *Journal of the Royal Society Interface*, vol. 10, no. 84, 2013.

[21] L. Eboli and G. Mazzulla, "Performance indicators for an objective measure of public transport service quality," *European Transport*, vol. 51, 2012.

[22] Direction générale de la mobilité et des routes (DGMR) du Canton de Vaud, "Rapport de l'observatoire de la mobilité," Direction générale de la mobilité et des routes (DGMR) Canton de Vaud, Lausanne, Switz, Tech. Rep., 2015.

[23] Bundesamt für Raumentwicklung (ARE), "Erstellung des nationalen personenverkehrsmodells für den öffentlichen und privaten verkehr - modellbeschreibung," Bundesamt für Raumentwicklung (ARE), Bern, Switz, Tech. Rep., 2006.

[24] M. Friman, B. Edvardsson, and T. Gärling, "Frequency of negative critical incidents and satisfaction with public transport services," *Journal of Retailing and Consumer Services*, vol. 8, pp. 95–104, 2001.

[25] C. Parkan, "Measuring the operational performance of a public transit company," *International Journal of Operations & Production Management*, vol. 22, no. 6, pp. 693–720, 2002.

[26] S. Chien and Z. Qin, "Optimization of bus stop locations for improving transit accessibility," *Transportation Planning and Technology*, vol. 27, no. 3, pp. 211–227, 2007.

[27] A. El-Geniedy, M. Grimsrud, R. Wasfi *et al.*, "New evidence on walking distances to transit stops," *Transportation*, vol. 41, pp. 193–210, 2014.

[28] A. Peungnumsai, H. Miyazaki, and A. Witayangkurn, "A grid-based spatial analysis for detecting supply-demand gaps of public transports: a case study of the bangkok metropolitan region," *Sustainability*, vol. 12, no. 10382, 2020.

[29] K. Kaeoruean, S. Phithakkitnukoon, M. Demissie *et al.*, "Analysis of demand-supply gaps in public transit systems based on census and gtfs data: a case study of calgary, canada," *Public Transport*, vol. 12, pp. 483–516, 2020.

[30] M. Schultheiss, "Assessment of the bus transit network: a perspective from the daily activity-travel organization of travelers," *Sustainability*, vol. 14, no. 2406, 2022.

[31] G. Drevon and A. Gumy, "Understanding multimodality through rhythm of life," in *Handbook of Urban Mobilities*. London, UK: Routledge, 2020.

[32] G. Dupuy, *L'urbanisme des réseaux : théories et méthodes*. Paris: Armand Colin, 1991.

[33] État de Vaud, "Projet d'agglomeration Lausanne-Morges (PALM)," https://www.vd.ch/themes/territoire-et-construction/amenagement-du-territoire/projets-dagglomeration/lausanne-morges-palm, accessed 21 Dec. 2023.

[34] E. Frejinger, M. Bierlaire, J. Stojanovic, M. Vrtic, N. Schüssler, and K. W. Axhausen, "A route choice model in Switzerland based on RP and SP Data," *Arbeitsberichte Verkehrs- und Raumplanung*, vol. 374, 2006.

[35] A. Meister, M. Felder, B. Schmid, and K. W. Axhausen, "Route choice modeling for cyclists on urban networks," *Transportation Research Part A: Policy and Practice*, vol. 173, 2023.

[36] Bundesamt für Energie, "Kennzahlen Fahrzeuge," https://www.bfe.admin.ch/bfe/de/home/versorgung/statistik-und-geodaten/kennzahlen-fahrzeuge.html, accessed: 2024-04-29.

[37] Mobilis Vaud, "Tarifs," https://www.mobilis-vaud.ch/fr/tarifs/, accessed: 2024-04-29.

[38] M. Bierlaire, "A short introduction to Biogeme," Transport and Mobility Laboratory, ENAC, EPFL, Tech. Rep. TRANSP-OR 230620, 2023.

[39] B. Schmid, F. Becker, K. W. Axhausen, P. Widmer, and P. Stein, "A simultaneous model of residential location, mobility tool ownership and mode choice using latent variables," *Transportation Research Part A*, no. 178, 2023.

[40] L. de Grange, P. Fariña, and J. Ortúzar, "Dealing with collinearity in travel time valuation," *Transportmetrica A: Transport Science*, vol. 11, no. 4, pp. 337–332, 2014.

Appendix

A.1 Additional public transport context

Name	Description
Car-exclusive	Activities structured entirely around the car. Car use is exclusive.
Private motor transport predisposed	Prefer the freedom in space and time, attached to quick travel, using multiple modes of transport.
Efficiency comparators	Most strongly persuaded by the efficiency of the transport modes; they will take whatever is fastest and most economical.
Comfort comparators	Search for comfort and time-saving; prefer to use the journey as free time.
Individual mode predisposed	Autonomy-motivated, avoid collective transportation as much as possible.
Alternative mode predisposed	Do not like to drive, specifically because of the stress induced while doing so.
Active mode predisposed	Avoid motorized modes for reasons of autonomy, physical activity, or reconnection with surroundings.
Environmentalists	Prioritize eco-friendly transport more than anything; image of different modes strongly influenced by their values.

Table A-1: Mode typologies used in comparison of mode choice logic, from [4].

(a) Geneva (Source: tpg)

(b) Lausanne (Source: t-l)

(c) Bern (Source: BERNMOBIL)

(d) Yverdon (Source: Travys)

(e) Nyon (Source: TPN)

Figure A-1: Public transport networks in the five selected communes.

Service aspect	**Indicator(s)**
Availability (sometimes accessibility)	Additional travel time relative to car (route directness), spacing distance between adjacent routings (coverage), stop spacing, density of lines, **minimum walking distance**, service frequency
Reliability	**Percentage of on-time arrivals/ departures**, headway regularity, running time adherence
Comfort	Degree of crowding relative to vehicle capacity (supply/demand gap)
Cleanliness	Cleaning frequency
Safety	Probability of road accidents, number of in-trip passenger fatalities
Fare	**Cost of one-way ticket**, cost of transfer, availability and volume of discounts
Information	Provision of supportive/confirming in-trip information, availability of functioning information devices on-board and at stops
Customer care	Extent and ease of ticket-selling network, ratio of employees in uniform
Environmental impacts	Number of ecological vehicles, noise and **pollutant emissions**

Table A-2: Performance indicators for various service aspects (after Eboli, 2012) [21].

A.2 SRHI Questionnaire

Behavior X is something...	
1.	I do frequently.
2.	I do automatically.
3.	I do without having to consciously remember.
4.	that makes me feel weird if I do not do it.
5.	I do without thinking.
6.	that would require effort not to do it.
7.	that belongs to my (daily, weekly, monthly) routine.
8.	I start doing before I realize I'm doing it.
9.	I would find hard not to do.
10.	I have no need to think about doing.
11.	that's typically "me."
12.	I have been doing for a long time

Table A-3: Full SRHI questionnaire, from [17].

A.3 Hidden Markov Model

For an HMM assumption to be valid, the hidden state must be reducible to a stationary first-order Markov process. In this case, the assumption of trip category as the hidden state implies that an individual's distribution of trip categories and mobility motifs do not change significantly over the observation window. This is in line with [30], who found that even after perturbations, strongly habitual behavior and high-level mobility patterns (excluding mode choice) persist, suggesting that habitual behavior is stationary. This also makes sense with intuition, as life circumstances such as employment status, income, and car ownership driving long-term habitual mobility behavior are not likely to significantly change over the course of a few weeks of observation. Regarding the reduction to a first-order Markov process, this is also reasonable, as we have seen that the choice of next destination is conditional upon the current location, with a degree of stochasticity: assuming a stationary process, this can be fully captured by a matrix of transition probabilities.

The main advantage of using HMMs is that the transition probabilities between observed variables (i.e. successive mode choices) can be decoupled from a hidden variable (i.e. trip purpose) whose transition probabilities are better-defined. This is a reasonable assumption in the case of mode choice habits [18], as we have seen that repeated places, purposes, and trip motifs are the strongest aspects anchored to habits, regardless of the mode chosen previously. The equilibrium emission probabilities are then estimated using numerical simulation given an initial transition probability matrix A.

Given a sequence of mode choices $S_M = \{\text{car}, \text{car}, \text{walk}, ..., \text{bus}\}$ as the observed variable and a sequence of category choices $S_C = \{\text{work}, \text{home}, \text{shop}, ..., \text{social}\}$ as the hidden state variable, an initial state transition matrix A is estimated from the observed transition frequency between each pair of states:

$$A = \begin{bmatrix} a_{(c_1,c_1)} & a_{(c_1,c_2)} & \cdots & a_{(c_1,c_{n_C})} \\ a_{(c_2,c_1)} & a_{(c_2,c_2)} & \cdots & a_{(c_2,c_{n_C})} \\ \vdots & \vdots & \ddots & \\ a_{(c_{n_C},c_1)} & a_{(c_{n_C},c_2)} & & a_{(c_{n_C},c_{n_C})} \end{bmatrix} ; \quad a_{(i,j)} = \frac{\text{trips}_{(i,j)}}{\sum_i \text{trips}_i}$$

where $\text{trips}_{(i,j)}$ is the observed number of trips made from state i to state j, and n_C is the total number of categories in habitual context C. The simulation is run to obtain the emission matrix (or, "equilibrium conditional probability matrix") B:

$$B = \begin{bmatrix} P(m_1|c_1) & P(m_1|c_2) & \cdots & P(m_1|c_{n_C}) \\ P(m_2|c_1) & P(m_2|c_2) & \cdots & P(m_2|c_{n_C}) \\ \vdots & \vdots & \ddots & \\ P(m_{n_M}|c_1) & P(m_{n_M}|c_2) & & P(m_{n_M}|c_{n_C}) \end{bmatrix}$$

which describes, for each category c and mode m, the conditional probability P of observing mode m given category c.

The HMM approach is incredibly flexible. An HMM approach may be generalized to any trip characteristic, including departure time, destination importance, trip length, trip duration, etc. It may even be extended to capture multiple characteristics at once, resulting in a chain of nested Markov processes. It can even be further extended by relaxing the first-order assumption; in fact, in further research, this may prove incredibly fruitful in habit quantification, as the current first-order assumption does not capture any effect of previous mode choice influencing the next, but rather that the mode choice is fully determined by the current state. Future research could couple these two, incorporating both the influence of the previous mode choice as well as the current state on the mode chosen in the future step.

A.4 Parking costs

Parking duration (hr)	Parking rate (CHF/hr)
0-1	4.00
1-2	3.50
2-8	3.00
8-10	1.50
10-12	0.00

Table A-4: Parking cost schedule used for parking cost calculation.

Parking costs were assigned according to the schedule above. For trips ending in the city center (and if the owner owned a city parking permit), parking cost was set to 0. If the destination was home or work, and free parking was available at home and work locations, respectively, then the parking cost was also set to 0.

A.5 Full model results

A.5.1 Utility equations

$$
\begin{aligned}
U_{\text{car}} = \text{ASC}_{\text{car}} & \\
& + \beta_{\text{time, car}} \cdot \text{BoxCox}(\text{time}_{\text{car}}) \\
& + \beta_{\text{work parking, car}} \cdot \text{work parking} \cdot \text{purpose}_{\text{work}} \\
& + \beta_{\text{cost}_{\text{fuel}}} \cdot \text{cost}_{\text{fuel}} \\
& + \beta_{\text{cost}_{\text{parking}}} \cdot \text{cost}_{\text{parking}} \\
& + \beta_{\text{income–cost}_{\text{fuel}}} \cdot \frac{\text{income}}{\text{mean income}} \cdot \text{cost}_{\text{fuel}} \\
& + \beta_{\text{income–cost}_{\text{parking}}} \cdot \frac{\text{income}}{\text{mean income}} \cdot \text{cost}_{\text{parking}} \\
U_{\text{pt}} = \text{ASC}_{\text{PT}} & \\
& + \beta_{\text{time access}_{\text{PT}}} \cdot \text{BoxCox}(\text{time access}_{\text{PT}}) \\
& + \beta_{\text{time egress}_{\text{PT}}} \cdot \text{BoxCox}(\text{time egress}_{\text{PT}}) \\
& + \beta_{\text{time wait}_{\text{PT}}} \cdot \text{BoxCox}(\text{time wait}_{\text{PT}}) \\
& + \beta_{\text{time invehicle}_{\text{PT}}} \cdot \text{BoxCox}(\text{time invehicle}_{\text{PT}}) \\
& + \beta_{\text{directness}_{\text{PT}}} \cdot \text{directness}_{\text{PT}} \\
& + \beta_{\text{access headway}_{\text{PT}},\ \text{PT}} \cdot \text{access headway}_{\text{PT}} \\
& + \beta_{\text{egress headway}_{\text{PT}},\ \text{PT}} \cdot \text{egress headway}_{\text{PT}}
\end{aligned}
\tag{A.1}
$$

$$
\begin{aligned}
&+ \beta_{\text{transfers}_{\text{PT}}} \cdot \text{transfers} \\
&+ \beta_{\text{cost}_{\text{PT}}} \cdot \text{cost}_{\text{PT}} \cdot (1 - \text{GA}) \cdot (1 - \text{demitarif}) \\
&+ \beta_{\text{cost, PT -- demitarif}} \cdot \text{cost}_{\text{PT}} \cdot \text{demitarif} \\
&+ \beta_{\text{income--cost}_{\text{PT}}} \cdot \frac{\text{income}}{\text{mean income}} \cdot \text{cost}_{\text{PT}} \qquad \text{(A.2)}
\end{aligned}
$$

$$
\text{where, directness}_{\text{PT}} = \left(1 - \frac{\text{time access}_{\text{PT}} + \text{time egress}_{\text{PT}}}{\text{time}_{\text{PT}}}\right) \;;
$$

$$
\begin{aligned}
U_{\text{bike}} = \beta_{\text{asc, bike}} &\\
&+ \beta_{\text{time, bike}} \cdot \text{BoxCox}(\text{time}_{\text{bike}}) \\
&+ \beta_{\text{average pos. slope}_{\text{bike}}\text{, bike}} \cdot \text{average pos. slope}_{\text{bike}} \\
&+ \beta_{\text{average slope}_{\text{bike}}\text{, bike}} \cdot \text{average slope}_{\text{bike}} \\
&+ \beta_{\text{infra. quality}_{\text{bike}}\text{, bike}} \cdot \text{infra. quality}_{\text{bike}} \qquad \text{(A.3)}
\end{aligned}
$$

$$
U_{\text{walk}} = \beta_{\text{time}_{\text{walk}}} \cdot \text{BoxCox}(\text{time}_{\text{walk}}) \qquad \text{(A.4)}
$$

A.5.2 Parameter estimates

Parameter	**Value**	**Rob. t-test**
ASC_{bike}	-8.419***	-15.698
ASC_{car}	-7.397***	-14.582
ASC_{PT}	-8.103***	-14.679
$\beta_{\text{access headway, PT}}$	0.00	-0.017
$\beta_{\text{ave. pos. slope, bike}}$	0.059***	3.621
$\beta_{\text{ave. slope, bike}}$	-0.038*	-2.425
$\beta_{\text{cost fuel}}$	0.699**	2.707
$\beta_{\text{cost parking}}$	-0.009	-1.543
$\beta_{\text{cost PT}}$	-0.201**	-3.237
$\beta_{\text{cost PT–demitarif}}$	-0.387**	-3.061
$\beta_{\text{egress headway, PT}}$	-0.001	-1.014
$\beta_{\text{income–cost fuel}}$	0.058	0.408
$\beta_{\text{income–cost parking}}$	-0.009	-1.395
$\beta_{\text{income–cost PT}}$	-0.113*	-2.436
$\beta_{\text{percent bike infra., bike}}$	0.002	0.385
$\beta_{\text{PT directness–time, PT}}$	8.458***	16.008
$\beta_{\text{time access, PT}}$	-0.307**	-2.957
$\beta_{\text{time, bike}}$	-0.015*	-2.394
$\beta_{\text{time, car}}$	0.540***	3.842
$\beta_{\text{time egress, PT}}$	0.00	-0.255
$\beta_{\text{time in-vehicle, PT}}$	-1.674***	-4.295
$\beta_{\text{time wait, PT}}$	-0.001	-0.481
$\beta_{\text{time, walk}}$	-2.677***	-6.326
$\beta_{\text{transfers, pt}}$	-1.06***	-4.275
$\beta_{\text{work parking, car}}$	-0.019	-0.055
$\lambda_{\text{BoxCox(time access), PT}}$	-0.084*	-2.678
$\lambda_{\text{BoxCox(time), car}}$	-0.898***	-6.282
$\lambda_{\text{BoxCox(time egress), PT}}$	4.281	1.309
$\lambda_{\text{BoxCox(time in-vehicle), PT}}$	-0.060***	-8.842
$\lambda_{\text{BoxCox(time wait), PT}}$	2.732**	2.874
$\lambda_{\text{BoxCox(time), walk}}$	-0.151***	-20.059

Table A-5: Full parameter estimates.

where * is significant at $p < 0.10$, ** is significant at $p < 0.01$, and *** is significant at $p < 0.001$. PT directness is the fraction of total PT time spent in-vehicle.