F. Verheij

Determinants and development of multimodal travel patterns

Identifying travel user groups in The Netherlands using Latent Class Cluster Analysis

TUDelft

Determinants and development of multimodal travel patterns

Identifying travel user groups in The Netherlands using Latent Class Cluster Analysis

by

F. Verheij

Master thesis submitted to Delft University of Technology
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE
in Complex Systems Engineering and Management
at the Faculty of Technology, Policy and Management,

to be defended publicly on Wednesday August 23, 2023 at 11:15 AM.

Floor Verheij
Student number: 4698398

Graduation Committee:

First Supervisor:	Dr. ir. M. Kroesen,	TU Delft, Section Transport and Logistics
Second Supervisor:	Dr. N. Goyal,	TU Delft, Section Organisation and Governance
Chairperson:	Dr. ir. M. Kroesen,	TU Delft, Section Transport and Logistics

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft

Preface

Before you lies the master thesis "Determinants and development of multimodal travel patterns: Identifying travel user groups in The Netherlands using Latent Class Cluster Analysis", which marks the end of my studies. This thesis has been written for the course CoSEM Master Thesis (SEN233) to fulfil the partial graduation requirements of the 2-year MSc Complex Systems Engineering & Management programme at the Delft University of Technology, Faculty of Technology, Policy and Management, The Netherlands.

For about half a year, I have worked full-time on this piece in the field of travel behaviour research. Within our master, we had the opportunity of specialising in one of the tracks offered, besides the other courses in the curriculum. I have followed the Transport and Logistics track, likewise, the domain I have followed in the three-year BSc Systems Engineering, Policy Analysis & Management (in Dutch: Technische Bestuurskunde) at our faculty. I am glad to have worked in this specific, challenging and exciting research field in the transport track and by doing quantitative analysis with 'LCA', one of the compelling topics and methods, among various others, which we have learned during our student journey.

Besides the inspiring environment our faculty offered, for which I am grateful, I would like to thank my supervisors Dr. ir. Maarten Kroesen and Dr. Nihit Goyal for their guidance, patience and helpful remarks. Finally, I would like to thank my family for being there for me during the graduation period, which felt like a never-ending process.

Floor Verheij

August 2023

Abstract

Passenger traffic by car is regarded as one of the main contributors to energy consumption and emission in the transport sector. Car dependency and limited shifts to more carbon-friendly alternative travel modes in industrialised countries play a major role in maintaining unsustainable mobility systems, despite governments' increased attention and effort in enhancing multimodal travel behaviour. In other words, having a diverse mode usage as a travel user. Compared to travellers who only use the car, multimodal travellers are likely easier to shift to more sustainable and health-enhancing modes, such as the bicycle and public transport, when applying policies. However, how (multi)modal travel patterns developed over time and the determinants of being multimodal are not often researched in combination with measuring multimodality and showing corresponding travel patterns. A Latent Class Cluster Analysis is performed once for 2010-2017 using cross-sectional data from the Dutch National Travel Survey (OViN) to measure multimodality whilst capturing distinct travel user groups per year, based on the frequency of travel mode use. Socio-demographic, mobility resource and built-environment variables are included as potential determinants of belonging to a specific group. The main results are that overall mobility patterns of the captured travel user classes were hardly subject to change. Moreover, only the smallest two out of five identified classes have travel behaviour with a higher degree of multimodality. Besides, the likely strong effect of owning mobility resources or not (e.g., a licensure, household car, company car, or household bicycle) on being likely in a car-dependent or a multimodal travel user group is shown. Most remarkably, our findings add to the existing knowledge by revealing that company car ownership plays a significant role in being a car-dependent travel user. Based on our results, identified policy directions include, but are not limited to, affecting mode choices of employees (owning a company car) via employer-based programs to incentivise them to use the bicycle or public transportation. Nevertheless, the knowledge about several travel user classes comprising multimodal travel patterns can be extended in several areas, most notably, by including attitudinal factors which could be tracked longitudinally on the individual level, such as perceptions about willingness to use travel modes, to acquire a more profound view about what strives people to behave in a certain way over a more extended period.

Keywords: Multimodality, Mobility patterns, Trends, Travel behaviour, Latent class cluster analysis

Executive summary

The effects of global warming are becoming more severe, like weather or climate extremes, and are impacting our lives. One of the main contributors to global warming are the emitted greenhouse gas emissions by human activities. Especially the transport sector has a high cause of this. Regarding passenger traffic, road traffic by car is the main contributor to high energy consumption and emission. A general rise in ownership of private cars and company cars, which are made available to employees by the company, is visible in countries' mobility systems. Car dependency and limited shifts to more carbon-friendly alternative travel modes in industrialised countries played a major role in maintaining unsustainable mobility systems, despite adopted governmental plans because of increased environmental awareness. Researchers proposed that increasing the multimodal travel behaviour of travel users (i.e. a diverse mode usage) is a potential way to stimulate more sustainable mode usage, implicating social relevancy such as limiting environmental impacts and improving public health and overall prosperity. However, multimodal travel behaviour is only limited practised.

Multimodality is in literature defined as a diverse (and balanced) mode usage, which means that someone has the highest level of multimodality when a variety of modalities are used, with about equal intensities, in a specific time period. Previous studies found that multimodal travellers are more sustainable than monomodal car users. Interestingly, multimodal travellers are also more likely to change their behaviour to more active or sustainable modes when the right conditions are provided. Within multimodal behaviour, the use of active travel modes (requiring a physical effort) like walking and cycling has a unique role. Besides the carbon- and health-friendliness of these modes, it is argued that a strict car user is less likely to switch to public transport, in contrast to a car user who already occasionally uses a bicycle, for instance, as it can be easily used as an access and egress mode for public transport.

Several aspects remain to be explored in the interest of researching multimodal travel behaviour. First, the literature has not yet agreed upon how to measure multimodality. Second, socio-demographic, (including household-related), travel mode availability and urban context determinants of multimodal behaviour showed some critical factors in previous research. Despite the importance of including mobility resources in research and the known car dependency, hardly any study involved company car ownership as a determinant in multimodal behaviour research. Third, studies investigating trends in measured multimodality are scarce, but, more importantly, the development of corresponding modal travel patterns, including various combinations of travel mode usage, remains to be determined.

Latent Class Cluster Analysis (LCCA), also named Latent Class Analysis (LCA), allows for measuring multimodality whilst enabling capturing modal travel patterns themselves. In short, LCA can capture comparable travel patterns of individual travel users and cluster them into homogenous groups, which are unobservable in real life and can be emergent and specifically targeted. Methodology-wise, determinants and a time variable in LCA are included to analyse the developing size of the classes reflecting (multi)modal travel patterns and the effects of determinants on belonging to one of the classes as a travel user. The identified knowledge gaps and the suitable method of analysing comparable groups resulted in the following main research question (MRQ): *"How are travel user groups comprising multimodal travel patterns characterised, and how do determinants and time influence the travel behaviour?"* to analyse the determinants and development of characterised multimodal travel patterns.

The multimodal travel patterns were once identified for individual adult travel users doing daily travel in The Netherlands from 2010-2017. High-quality yearly cross-sectional national data from the Dutch National Travel Survey (OViN) is used, in which individual participants filled in a one-day travel diary about the trips they made, including their personal and household characteristics. The modalities walking, cycling, public transport (bus/tram/metro/train), and car (driver or passenger) are used to define travel patterns. The level of analysis is at the stage level, as it accounts for the variation in all possible modes used in a time period. Stages are part of a trip, e.g., a trip can comprise a walking, public transport, and cycling stage. So, the number of stages per travel mode is used as a measure of intensity for mode usage. The degree of multimodality of the travel user classes is analysed for the multimodality measurement number of modes and OM_PI (Objective Mobility Personal Index), an indicator measuring the diversity (of mode usage) and equality (balanced mode use), to provide a simple one-sided view and a complex multi-sided view, based on our literature study. Moreover, socio-demographic, mobility resource and built-environment variables are included as determinants based on our literature review.

In our results, we captured the travel users of society into five distinct classes: a car and bicycle multimodal class (also involving some walking), a public transport plus multimodal class (involving also cycling, walking and limited car use), a car exclusive class, a car and walk class and a bicycle mostly class. The sample's most prominent classes (biggest) are the car exclusive and the bicycle mostly class. However, these classes showed the lowest multimodality measurements. On the other hand, the classes named multimodal have the lowest class sizes, whereas the public transport plus class is the smallest and the most multimodal. Several main findings are established and subsequently elaborated.

It is found that in our study, more multimodal behaviour is in general prevalent among females, individuals from immigrant origin, younger age groups and students, some higher educated individuals, individuals with lower household incomes, households with fewer members, those not having a car licensure, not owning a household car or company car and when owning a household bicycle. Mobility resource ownership variables are likely strong to determine latent class membership. Most remarkably of all findings, the findings for ownership of company cars add to the existing knowledge.

Furthermore, we have indicated that most classes are equally (but less precise) spread among the residential municipalities. However, for the public transport plus class, we have revealed a clear visible pattern in having a higher likelihood as a travel user residing in the urbanised Western part compared to the peripheral part, comparable to the finding of individuals with a higher urban density being more associated with multimodality.

Last, we have uniquely shown the development of multimodal travel patterns. The class sizes of the identified travel user groups barely changed, which remarks that the multimodal travel patterns of society, captured by the classes, have not developed over time. Given changes in demographics, fluctuating sample representativeness, and the increased awareness of the impact of climate change and active travel among individuals and practitioners, it has not accomplished the desired effects of creating more active, sustainable (and multimodal) travellers.

To conclude, mobility patterns are hardly subject to change, despite improved awareness of environmental impacts and having sustainability higher in the political agenda. This indicates room for improvement in policy-making. For instance, certain travel user groups can be targeted differently by governments to promote the use of multimodal hubs (physical places to enable switching modes) and the use of MaaS (Mobility as a Service, an app which helps to plan trips for travel users). Moreover, one of the other policy directions is employing a special role for employers to target (company) car-dependent travel users by providing employer-based programs (e.g., providing higher monetary compensation for bicycles) to attract employees using active travel modes or public transportation. Nevertheless, the knowledge about several travel user classes comprising multimodal travel patterns can be extended in several areas, most notably by including attitudinal factors which could be tracked longitudinally on the individual level, such as perceptions about willingness to use travel modes, to acquire a more profound view about what strives people to behave in a certain way over a more extended period.

Content

1. Introduction

1.1. Background on the environment, transport, and passenger multimodality

The greenhouse gas effect causes global warming, which is the rise of the earth's surface and sea temperatures. The effects of global warming, including weather and climate extremes, are becoming more severe and have an impact on our lives, according to the World Meteorological Organization (WMO, 2023). Nevertheless, global warming is mainly driven by increasing emitted greenhouse gases (GHG) from human activities (IPCC, 2022; WMO, 2023).

One of these activities is transport, i.e. travelling and transporting goods. According to the International Energy Agency (IEA), the global annual growth rate of transport emissions is nearly 1.7% from 1990 to 2021, higher than any other sector (IEA, 2022). In 2010, the global transport sector contributed to a quarter of all energy-related direct CO_2 emissions, as stated by the Intergovernmental Panel on Climate Change (IPCC, 2014). Despite adopted transport-related policies in countries and more efficient (passenger) vehicles, the GHG emissions continued to grow in the last decades (Brand et al., 2013; IPCC, 2014; Schafer & Victor, 1999). Also, after a historical traffic drop in 2020 influenced by national lockdowns and homeworking measures due to the Covid-19 pandemic, the emission growth by transport rebounded in 2021 (EEA, n.d.; IEA, n.d.). In 2021, the IEA estimated that the transport sector accounted for 37% of CO_2 emissions among end-use sectors (IEA, n.d.).

When considering the transport sector's main travel modes for passengers and freight, many industrialised countries have dominated car use for the past decades (Gerlofs-Nijland et al., 2021; Olde Kalter et al., 2020). The energy consumption from road traffic by passenger cars specifically was almost 60% in 2018 in IEA economies (Elghozi, 2021). The high usage and pollution of passenger cars have multiple reasons, like the high reliance on fossil fuels and the limited mode replacement to more carbon-friendly alternative modes (IEA, 2022).

In the interest of decarbonising passenger travel and limiting the car-dependency, multimodal travel behaviour has reached more attention among researchers. Passenger multimodality is generally defined as having a diverse mode usage as a travel user (e.g., Nobis, 2007). It is found that multimodal travellers are generally more sustainable (emitting less CO_2) than monomodal car users when both trips involve similar distances (Heinen & Mattioli, 2019b; Nobis, 2007). Despite this, multimodal behaviour does not necessarily result in less car use (An et al., 2021) or sustainability per se, as car-dependant multimodal groups also exist (Hunecke et al., 2020). On the other hand, when only using one mode or two modes from the mode set bicycle and walk and thus having monomodal or bimodal behaviour, it is more sustainable than multimodal behaviour involving non-active modes, which require no physical effort.

Nonetheless, studies found that (for short-distance trips) multimodal travellers are more likely to change their behaviour to more active or sustainable modes after policy interventions in certain conditions (Heinen & Ogilvie, 2016; Prato et al., 2017). Götschi et al. (2017) explain that consumers (i.e. travellers), in general, weigh the (dis)advantages of current options. Depending on the situation, another mode might be preferred for one trip or another, which can indicate a swift. Furthermore, Heinen (2018) explicates that research suggests that more multimodal travellers are more likely to switch modes.

Within all modes, cycling (active travel) can be seen as an intermediary mode related to multimodal behaviour. Cycling intermediates multimodal behaviour because the bicycle can be combined with- or easily switched to another mode when cycling is unsuitable (Kuhnimhof et al., 2010; Olafsson et al., 2016; WHO, 2022). It is also explained by Kroesen (2014) that a strict car user (only using the car) is less likely to switch to public transport (PT), in contrast with a car user who already occasionally uses a bicycle, because the bicycle can play as an access and egress mode for public transport, which involves using multiple modes.

Multimodal behaviour in the transport sector is also regarded among practitioners as a possible way to stimulate more sustainable and active travel among travel users and enhance societal impacts like less environmental harm and improved health, resulting in more awareness about this topic in the last few years. From a governmental perspective, countries are, from 2018 onwards, more focused on integrating active mobility, which requires physical effort, and sustainable mobility in the transport system (European Commission, n.d.). This is to promote the variety of mode usage in trips. Moreover, a resolution to promote (walking and) cycling in the transport system is adopted by the 193 global members of the United Nations General Assembly (Klingert, 2022).

Empirically seen, multimodal behaviour is globally visible in practice (e.g., Gerlofs-Nijland et al., 2021; Klinger, 2017). For example, a new mobility mix generally arose because digital service innovations enabled multimodal travelling (WHO, 2022). However, the high car dependency is still visible in a risen private car ownership and company car (made available to employees by the company) ownership. An example, in The Netherlands, car ownership is growing faster than the population (Statistics Netherlands, 2020a). Moreover, the Dutch National Institute for Public Health and the Environment (RIVM) argues that the current global transport system and mobility patterns involving modal travel patterns remain unsustainable and that multimodal travelling is mainly visible in cities or practised by the younger generation using smartphones, such as for sharing services (Gerlofs-Nijland et al., 2021).

1.2. Theoretical background, previous studies and knowledge gaps on multimodal travel behaviour

Theoretical background

As the notion of passenger multimodality (multimodal travel behaviour) and the problem situation is introduced (section 1.1), a theoretical background is given to gain an understanding of the research stream about multimodal travel behaviour. Moreover, multimodality and measurements of multimodality are defined to find knowledge gaps in the body of literature for formulating research objectives (section 1.5).

To date, much research focussed on interpersonal variability, aiming at explaining differences (the variety) in travel behaviour between individuals based on personal or context-related attributes. Whereas little research has been done about a more recent stream, the intrapersonal variability (Faber et al., 2022; Heinen & Chatterjee, 2015; Scheiner et al., 2016; Zhang et al., 2021). Intrapersonal variability means having a variable travel pattern as an individual. Variability in travel patterns, circumstances, and preferences, includes but is not limited to trip purposes (e.g., commuting or leisure travel), activities, destinations, time patterns, weather conditions, available resources or mode choice (Heinen & Chatterjee, 2015; Scheiner et al., 2016). A variable travel pattern could thus mean for a person that the time patterns of their travel behaviour or the available resources vary. Zhang et al. (2021) argue that ignorance of the variability of individuals' travel behaviour from trip to trip, day to day or week to week may lead to bias in travel demand model estimations. These estimations can be used for policies.

Besides that individual (circumstances of) travel patterns may vary over some time, a repetition in travel behaviour generally occurs (Heinen & Chatterjee, 2015; Schlich & Axhausen, 2003). So, people tend to have repeated archetypical daily patterns over a week. Still, it is reasoned that it is likely that intrapersonal modal or mode choice variability, i.e. the variable use of transport modes by individuals, exists irrespectively of variability in travel patterns in the broadest sense. This implies that although a broader travel pattern of an individual (e.g., captured by time patterns or trip purposes) might vary or not, a variety of transport modes is likely to be used by an individual in a period of time. Within intrapersonal variability research, mode variability, as opposed to activity, destination, or time pattern variability, has achieved less attention yet (Heinen & Chatterjee, 2015). As with interpersonal variability, differences between individuals or groups can still be captured when intrapersonal variability is examined.

A higher modal variability means that, within the travel pattern of an individual, more different modes are used during a time period, and a person is thus regarded as more multimodal. Multimodality is thus generally defined as the (flexible) use of various modes during a specific time period (Buehler & Hamre, 2015; Kuhnimhof, Armoogum, et al., 2012; Nobis, 2007). A monomodal traveller is then the lowest possible degree of multimodal travel behaviour. In fact, by a strict definition, unimodal travellers are not multimodal. A higher degree of multimodality, for instance, is seen when a trip by an individual involves multiple stages, for instance, a cycling stage, a public transport stage and a walking stage, compared to someone using only the car (one stage) in a trip. The example if over the course of one trip, which is actually defined as intermodality, a subset of multimodality is intermodality, which is defined over the course of one trip. The general scope of multimodality is intrapersonal variability, which looks broader than one trip. This provides a broader view of individual multimodality (Heinen, 2018). More specifically, the actual travel pattern and multimodality level are better captured when looking at multiple trips. Because a daily, multiple-day or longer time period likely involves multiple trips.

Besides that the usage of various modes generally defines multimodality (throughout a time period or multiple trips), the body of literature emphasises that other requirements are essential for defining the degree of multimodality, like the modal intensity or the types of modes used (Nobis, 2007). A balanced mode use without predominance is mainly regarded as a high multimodality, next to the variable mode usage (Diana & Pirra, 2016).

Multimodality can be measured using various measurements, as explained in the literature review (part of desk research) in Chapter 2, and several measurements are used in the literature used for the review in Chapter 3. In short, no consensus has been reached yet about which multimodal measurements can be used for which application (e.g., An et al., 2021; Diana & Pirra, 2016; Heinen, 2018; Heinen & Mattioli, 2019a; Heinen & Chatterjee, 2015; Scheiner et al., 2016). The most prevalent measurements are predefined nominal groups, in which predefined requirements define to which predefined modal group someone belongs. For instance, based on the number of modes or combinations of mode usage, someone can be placed in a monomodal car user group when only the car is used. Moreover, data-driven measurements are widely used, which define post-hoc and not a-priori nominal groups. Depending on the data and specific characteristics of individuals, travel segments emerge based on clustering schemes. Other (quantitative) measurements are more intuitively used and 'one-sided' compared to multi-sided requirements, as the share of the primary mode could define how multimodal someone is. Last, continuous indicators involve both the mode use variability and modal intensity, captured in a metric, which is regarded as a 'multi-sided' view.

Multimodality (summary)

The (flexible) use of various travel modes during a specific time period.
A more balanced mode use is also often regarded as a higher degree of multimodality. Other requirements, such as the modal travel intensity or the types of modes used, can also be used to assess the degree of multimodality.

Previous studies

As the theoretical background is given, including background on the research stream on multimodal travel behaviour, the multimodality definition and measuring multimodality, previous research is discussed to identify more knowledge gaps for formulating research objectives (section 1.5).

The determinants of multimodality are extensively researched. Many found more multimodal behaviour among younger age groups, higher-educated ones and in urban settings (e.g., Hunecke et al., 2020; Lee et al., 2020; Nobis, 2007). Moreover, multimodal travellers are generally more prevalent among women, white ethnic groups, students, individuals with higher incomes, and households with fewer members, based on the literature review in Chapter 3. Regarding urban density, a geographical variable, some studies showed the spatial distribution of multimodality indices, in the United States mainland (Lee, 2022), the city of Lisbon (Lemonde et al., 2021a, 2021b) or parts of a city (Ren et al., 2022) to find concentrations of areas with similar values measuring multimodality. Similarly, higher multimodal measurements were overall visible in more dense urban areas or areas with more infrastructural diversity. Next to socio-demographic, household-related, and urban context determinants, most of the literature focussed on the availability and ownership of specific travel modes (e.g., Klinger, 2017). As explained by Klinger (2017), briefly seen, people who own a car are, in most cases showing unimodal car travel behaviour, and people owning a bicycle tend to combine modes more often. Besides, summarised from Chapter 3, people who do not hold a car license and people with limited car availability show more multimodal behaviour. Car dependency is thus considered a widely important factor.

Despite the known car dependency, the ownership of a company car has so far not often been researched yet in combination with multimodal travel patterns, to the best of our knowledge. Company cars are referred to as cars made available to the employee for work and private purposes. Often a fuel card is supplied, which enables it to fill up at lower or no cost. Despite, it is found that company car owners drive significantly more (Van Eenoo, Boussauw & Fransen, 2022) and are the least susceptible to changing modes (Curtis & Headicar, 1997). Moreover, a cluster analysis showed that some company car owners have a long home-work distance, while others have a shorter commuting distance and use the car mainly for private trips (Macharis & De Witte, 2012). One example of researching company cars in relation to multimodality showed that a presence of a company car is mainly visible in car-dominant groups in Belgium (Van Eenoo, Fransen & Boussauw, 2022). Company car ownership could thus be regarded as an essential notion of car dependency, limiting multimodal behaviour, yet an unexplored research field.

Among multimodal studies, many focus on static multimodal groups, i.e. defining once multimodal groups based on one or multiple years (e.g., An et al., 2023; De Haas et al., 2018; Haustein & Kroesen, 2022; Molin et al., 2016; Ton et al., 2020). Some research studied multimodality longitudinally or described the main body of literature about which travellers are more likely to change their behaviour over time, due to life changes, for instance (e.g., Haustein & Kroesen, 2022; Klinger, 2017; Kroesen & Van Cranenburgh, 2016; Scheiner et al., 2016). Predominantly, these studies investigate travel patterns with panel data at two points in time, with an intermediate period of one up to five years, or it is studied hypothetically. Studies which analyse the trends and development of multimodality over a more extended period are scarce, while some other (only descriptive) studies suggest that (younger) travellers in industrialised countries are becoming increasingly multimodal (Buehler & Hamre, 2015; Kuhnimhof et al., 2011, Kuhnimhof, Armoogum, et al., 2012; Kuhnimhof, Buehler, et al., 2012; Streit et al., 2015).

One example of one of the few studies doing a trend study is a study by Heinen and Mattioli (2019a) in England from 1995-2015, which used cross-sectional data from the National Travel Survey. They found that multimodal behaviour, measured in multimodal indicators, decreased. Moreover, a shift towards monomodal daily travel is found. Another contribution to studying multimodal trends has the same context; the timespan is 1995-2017 (An et al., 2021). The findings showed that based on the predicted values of the multimodality measurement for several years, a slight declining trend is seen. Although some fluctuations are found, the magnitude of the range of changes is considered small.

To conclude this paragraph, some studies indicate an increasing multimodal trend. In contrast, the trend studies that sought to fill in the gap of a lack of multimodal trend analyses found a slightly decreasing or stable trend in measured multimodality. However, foremost, in which respect the corresponding modal travel patterns changed accordingly is not assessed. This raises the question of whether a trend in (the measured) multimodal behaviour and the modal travel patterns themselves can be seen.

Knowledge gaps

Summarising the previous, multimodal travel behaviour studies have shown important insights already. However, some knowledge gaps remain.

1. Evidently, the empirical importance of objectively comparing multimodality with a scientific basis is being more recognised (e.g., Diana & Pirra, 2016). However, until now, limited attention has been paid to the different measurements (e.g., Heinen, 2018; Heinen & Mattioli, 2019a). The literature has not yet agreed on how multimodality should be measured depending on the research.
2. Socio-demographic, household-related, travel mode availability, and urban context determinants of multimodal behaviour showed some critical factors in research (e.g., see Chapter 3). Travel mode availability is considered essential for having or adopting multimodal travel patterns or not (e.g., Klinger, 2017). Car dependency due to car ownership is inhibiting more multimodal behaviour (e.g., see Chapter 3). Despite the importance of including mobility resource variables in research, hardly any study involved company car ownership as a determinant in multimodal behaviour research (except for, e.g., Van Eenoo, Fransen & Boussauw, 2022), which remains to be explored.
3. Studies investigating trends in measured multimodality are scarce and mainly about the measured multimodality (e.g., An et al., 2021; Heinen & Mattioli, 2019a), but, more importantly, the development of modal travel patterns, including various combinations of modal usage, remains to be determined. Previous studies have not allowed us to acquire a comprehensive view of multimodal trends.

1.3. Policy relevance of multimodal travel behaviour

Besides the scientific relevance of researching multimodal travel behaviour (section 1.2), multimodality is also highly relevant for policy-making, as some high-level policies of integrating modes in the mobility system to reach sustainability and limiting emission goals are mentioned earlier (section 1.1). The potential of encouraging multimodality to increase the sustainability of transport systems is visible in the following. Multimodal travellers, mainly involving active travel and generally more sustainable than car-dependent travel users, are likely easier to transition to sustainable transport modes (section 1.1, 1.2). The use of bicycles can also play a role towards a shift, as it is underpinned that it can intermediate multimodal behaviour in which public transportation is used. On the other hand, private car ownership and company car ownership exhibit car dependency, away from multimodal behaviour.

So, given that multimodality may be seen as a potential first step away from 'habitual' car dependency (Heinen & Mattioli, 2019a; An et al., 2022), policies promoting active modes and public transport, and positive attitudes towards them (An et al., 2022), or applying for employer-based programs in reducing company car usage (Pucher et al., 2010) could play an important role. This raises the question of how to stimulate multimodality and reduce car dependency effectively. Identifying determinants of modal travel patterns, like company car ownership, could be a point of departure for tailor-made policies. An important step is identifying (multi)modal travel patterns, in which some comparable modal patterns of individuals are more suited to incentives towards a modal shift. Moreover, for targeting, monitoring, and evaluating policies, tracing the development of travel patterns of comparable groups is essential (An et al., 2021, 2023; Ton et al., 2020). As Weller et al. (2020) explained, some comparable groups could benefit from interventions based on their shared characteristics. For instance, it can be used to design 'hard' or 'soft' policies. Namely harder infrastructural policies or softer policies for creating a better (information) environment and opportunities via management or marketing (Anable, 2005).

1.4. Societal relevance of multimodal travel behaviour

Next to the policy relevance (section 1.3), improving multimodality has societal impacts when active travel (walking and cycling) replaces car usage. Policies can achieve to limit environmental impacts (section 1.1), improving public health and overall prosperity by influencing consumers' mode choices (WHO, 2022; section 1.3). The environmental benefits of active travel are essential as the greenhouse gas effect increasingly affects our lives (section 1.1) and has become increasingly important for society over the years. In contrast to using active travel modes, physical inactivity is an important notion as it is one of the behavioural risk factors for morbidity and mortality worldwide (Lim et al., 2012). Furthermore, overall physical activity contributes to health, in which active travel can play an important role (e.g., Gerlofs-Nijland et al., 2021; Kroesen, 2014; Kroesen & Van Wee, 2022). Moreover, active travel is regarded as financially attractive. So, active travel within multimodal behaviour or modal shifts of multimodal travellers towards more active and sustainable modes can induce environmental, health and prosperity effects, according to the European Cyclists' Federation (ECF, 2018), which are now elaborated further.

Environmental benefits of active travel include little noise and air pollution (e.g., Buehler & Pucher, 2012; United Nations Economic Commission for Europe, 2021) and less space use of bicycles compared to cars (Buehler & Pucher, 2012), which is also the case for people who walk. Moreover, multimodal (or active) travellers generally make less use of non-renewable resources (with less CO_2) (e.g., Buehler & Pucher, 2012; section 1.1). For instance, 1 kilogram of CO_2 per 7 km travelled can be saved using a bicycle instead of a car (UNEP, 2019). Besides this, active travel can reduce congestion (European Parliament, 2023).

Within travel patterns, switching from car trips to low-carbon trips with modes like walking, cycling, and also public transport is considered relevant for health-enhancing policies (Maibach et al., 2009). Several (reviews of quantitative) studies found health benefits in several relationships between active travel and better health outcomes for people (e.g., Oja et al., 2011; Saunders et al., 2013; Wanner et al., 2012; Warburton, 2006). Health benefits include increased physical activity, fitness and well-being, and reduced disease risks (e.g., obesity).

Economic advantages due to active travel compared to cars for commuters and governments exist. The economic advantages for persons are because of a minimal personal financial investment in active travel modes leading to high accessibility for everyone (Buehler & Pucher, 2012). For instance, cycling can be easily integrated into day-to-day travel (Gerlofs-Nijland et al., 2021). Regarding governments, lower public infrastructure investments are required for active travel infrastructure (Buhler & Pucher, 2012).

1.5. Research objectives and main research question

Three knowledge gaps are identified in multimodal travel behaviour research. Nevertheless, researching this topic is highly relevant as such behaviour is limited empirically seen yet (section 1.1). Multimodality is also policy (section 1.3) and societal (section 1.4) relevant. The threefold research objectives are explained to fill in the identified gaps of several referred studies (section 1.2):

1. First, this thesis offers a comprehensive theoretical overview of multimodality measurements and their implications as a starting point to advance state-of-the-art on the different measurements. The choice of how empirical travel behavioural patterns with specific modalities can be measured objectively in our research is grounded in the field of research on multimodal travel behaviour.
2. Second, the socio-demographic, household-related and urban context (including geographical) determinants of multimodality are examined using empirical data. The inclusion of a broader set of mobility resource variables, meaning the availability and ownership of travel modes, including private car ownership and company car ownership, is put forward to broader investigate car dependency, an important factor in negatively hampering multimodal behaviour.
3. Third, this research takes a dynamic approach to multimodal travel behaviour. So, an emphasis is put on the longitudinal aspect to show the development of multimodal travel behaviour by looking at the modal travel patterns for a more extended period, apart from the measured behavioural trends of multimodal indicators in previous research, where researchers shed light on mainly.

   Latent Class Cluster Analysis (LCCA), hereafter named Latent Class Analysis (LCA), allows for measuring multimodality whilst enabling capturing modal travel patterns themselves by identifying travel groups consisting of individuals practising specific comparable patterns. When referring to multimodal patterns, it is about analysing multiple mode usages in travel patterns, and the degree of multimodality is about how multimodal those patterns are. The suitability of this quantitative method (compared to other methods) and the advance of once-identified travel user groups is further explained in Chapter 4 (section 4.1). Methodology-wise, a time variable in LCA is included to analyse the developing multimodal patterns. Somewhat comparable, the inclusion of several years is earlier seen in an LCA by De Haas et al. (2020) for identifying the development (group sizes) of comparable e-bike user groups in terms of comparable user characteristics. Other than that, the use of this methodological aspect of LCA is used to acquire a view of the development of groups involving modal travel patterns over several years, and determinants are added.

The following main research question (MRQ) is set up based on the previous research objectives:

MRQ	How are travel user groups comprising multimodal travel patterns characterised, and how do determinants and time influence the travel behaviour?

1.6. Research questions

Based on the previous research contributions, the main research question is derived (section 1.5). In order to characterise travel user groups comprising multimodal travel patterns and analyse the influences of determinants and time (MRQ), the research questions are formulated and listed in Table 1. First, insights on how multimodality is measured in previous studies are derived (RQ1), and insights of previous studies identifying multimodal travel behaviour determinants are derived (RQ2). Afterwards, trends of the descriptives of travel behaviour determinants and travel behaviour are analysed (RQ3). Besides, the identified travel user classes of multimodal travel patterns are explained, and their degree of multimodality is assessed (RQ4). The explanatory effect of potential determinants on class membership of travel users is also assessed (RQ5). Moreover, the spatial distribution of the class membership shows how multimodal travel patterns of travel users are distributed among municipalities to characterise the classes (RQ6). Finally, the development of travel user classes and corresponding (multi)modal travel patterns are shown over time by looking at the changing class sizes of the model outcome consisting of once-identified groups (RQ7).

Table 1. Research questions overview.

Number	Research question
MRQ	**How are travel user groups comprising multimodal travel patterns characterised, and how do determinants and time influence the travel behaviour?**
RQ1	Which insights on measuring multimodal travel behaviour can be derived from previous studies?
RQ2	Which insights on multimodal travel behaviour determinants can be derived from previous studies?
RQ3	What are the observed trends of travel behaviour determinants and travel behaviour?
RQ4	What are the captured multimodal travel patterns and degrees of multimodality by the identified travel user classes?
RQ5	To what extent do determinants influence class membership of travel users?
RQ6	What is the spatial distribution across municipalities of class membership of travel users?
RQ7	How do travel user class sizes and classes' multimodal travel patterns develop over time?

1.7. Research approach and scope

After the research is set and embedded in context in the previous sections, the broader research overview and research flow are given (section 1.7.1, research approach). After that, a summary of the scope of this research is given (section 1.7.2).

1.7.1 Research approach

The research approach is explained via the research overview and a research flow.

Research Overview

Figure 1 shows the research overview.

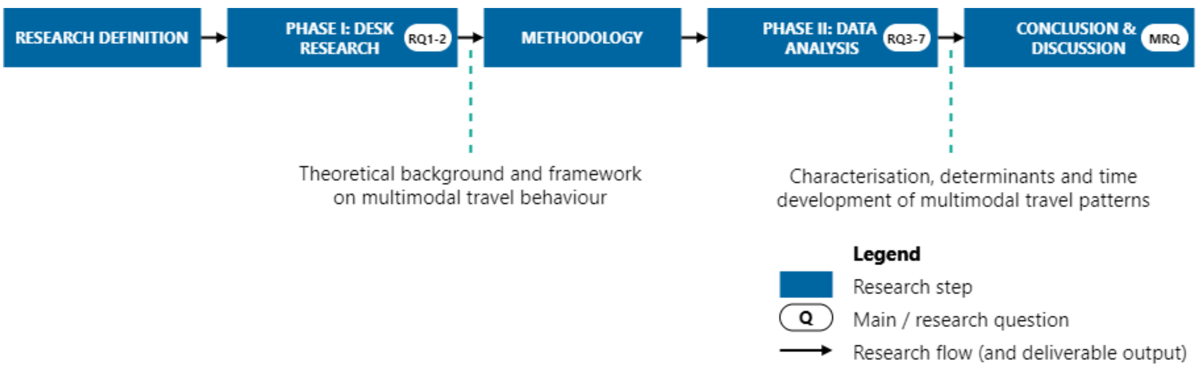


Figure 1. Research overview.

The research consists of two phases: desk research (1) and data analysis (2), and some additional research steps. In this current chapter, the research is already defined, and both phases will be shortly elaborated in this section. Every phase answers multiple research questions (defined in Table 1). The research methodology is provided before the main research phase (2), together with the (dis)advantages of this method and viewpoint used. So, the desk research phase (1) is subordinate to the data analysis phase (2), not meaning that it is less important, but the corresponding approaches (and suitability) are in that chapter itself elaborated. It means that the output 'theoretical background and framework on multimodal travel behaviour' of phase 1 (for RQ1-2) flows into the data analysis phase 2 (for RQ3-7), where the main outcomes of Latent Class Analysis will lead to the final deliverable output. The final deliverables are the 'determinants and development of travel user groups characterised by modal travel patterns'. The obtained knowledge leads to answering the main research question in the conclusion and discussion. A more detailed reading guide based on this research overview is in this chapter provided (section 1.9).

Research Flow Diagram

Figure 2 depicts the Research Flow Diagram. The two research phases, as shown in Figure 1, are visualised in this figure together with the research flow and (final) deliverable output. The research deliverables per research question (Table 1) and the internal research flows are given within each phase. For every research deliverable, the method and tools are also shown.

- **Phase I: Desk Research**

This first phase constitutes an overview and definition of multimodal travel behaviour measurements (RQ1). A literature study is done about the differences in measurements to decide which measurement to use. Moreover, an overview of multimodal travel behaviour determinants is made by systematically reviewing other literature (literature review) to compare potential explanatory determinants (RQ2). The methods and suitability aspects are explained in the corresponding chapters (see section 1.9). The definitions of multimodal travel measurements are internally flowing into this step, as it is important to know what multimodality is before analysing the determinants of it. The deliverable output, a theoretical background, is the basis for the conceptual framework on multimodal travel behaviour.

- **Phase II: Data analysis**

The second phase is about analysing and modelling the used survey data (section 4.3). The explanation of the research flow starts with RQ4. The multimodal travel patterns and degrees of multimodality in terms of modal usage and intensity of travel user classes, which emerged from using the LCA method for pooled data (section 4.1, including suitability), are analysed (RQ4). The travel user classes are thus characterised quantitatively and qualitatively (by classifying names), and whereafter, the information flow goes to the next step. This step is about the influence of determinants on class membership of travel users (RQ5). So, the explanatory effect of determinants on being a member of one of the identified travel user classes is assessed. Based on the defined model outcome (RQ4/5), the spatial distribution across the municipalities can be visualised to characterise the classes. It shows the distribution of class membership of travel users' residential municipalities for several classes (RQ6). After these outcomes, the previous model information (RQ4/5) can be used to assess the time component of the travel user classes because the class sizes differ per year (RQ7). So the development of corresponding multimodal travel patterns of the classes over time is provided. Moreover, the trend overview of descriptives of potential travel behaviour determinants and travel behaviour (RQ3) flows into this step to provide background information on class development. After all, this second phase delivers the final output of multimodal travel patterns' characterisation, determinants and time development.

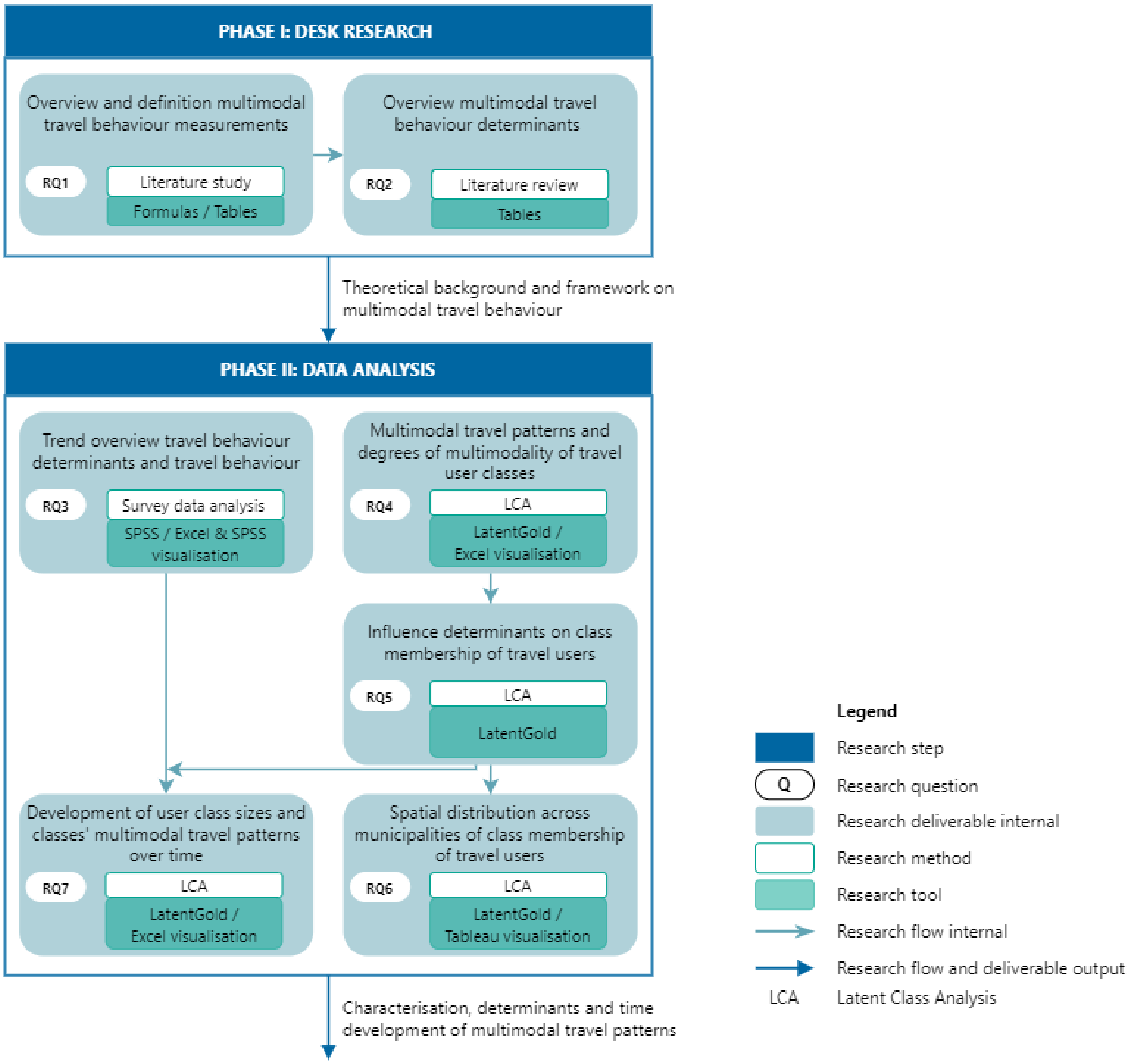


Figure 2. Research Flow Diagram.

1.7.2 Research scope overview

After the research approach is explained, the research scope is explained. The country of interest is The Netherlands. Many Western studies on multimodal travel behaviour are identified (Chapter 3), creating more comparability opportunities between results and opportunities to find potential determinants. Moreover, the data from our country is easy to obtain (section 4.3), making it more feasible to study. In this case, it is an active travel mode-oriented country, meaning that active travel modes, namely walking and cycling, are embedded in modal patterns (Kennisinstituut voor Mobiliteitsbeleid, 2019). Identifying modal travel patterns could thus lead to identifying various multimodal combinations, for which targeted policies can be created to enhance the use of active travel modes and multimodal travel behaviour further, making it worth studying. Although research has been done about multimodality in The Netherlands (Chapter 3), no research has been identified about (measuring) multimodal trends and company car usage. While still car dependency is an issue in The Netherlands, like other Western industrialised countries (section 1.1; Kennisinstituut voor Mobiliteitsbeleid, 2019), it contributes to making this a suitable exemplar to study.

1.8. CoSEM relevance

This defined research study (thesis) is relevant to the MSc programme Complex Systems Engineering and Management (CoSEM) since the use of travel modes by travel users part of a complex socio-technical system, which is the typical system of interest in this programme. Social aspects emerge because traffic participants in a (multi)modal trip interact in the traffic system by showing travel behaviour. Private parties are also involved, such as employers or the car industry. Employers are interested in healthy employees, but the car industry has competing interests than incentivising them to walk, use the bicycle or public transport. Combining these social and technical perspectives, the system is socio-technical given the technical traffic system consisting of designated mode or mixed modes infrastructure. The involved traffic users who dynamically change modes add to the complexity of the socio-technical system. Furthermore, more complexities are visible in the context of a traffic or mobility system, such as various politicians potentially persuading each other of traffic-related policies, which might stimulate or hamper higher multimodal travel usage.

The results of this study, which involves analysing similar mobility styles of user travel groups using the taught travel behavioural research method Latent Class Analysis, enable several insights for changing the socio-technical environment in which the modal travel patterns are part of, such as developing institutional or infrastructure-related policies that interfere with the design of the transportation systems in a country.

1.9. Reading guide

The remainder of this Master's thesis is as follows. The chapter overview is visualised in Figure 3, which is an addition to the research overview in Figure 2.

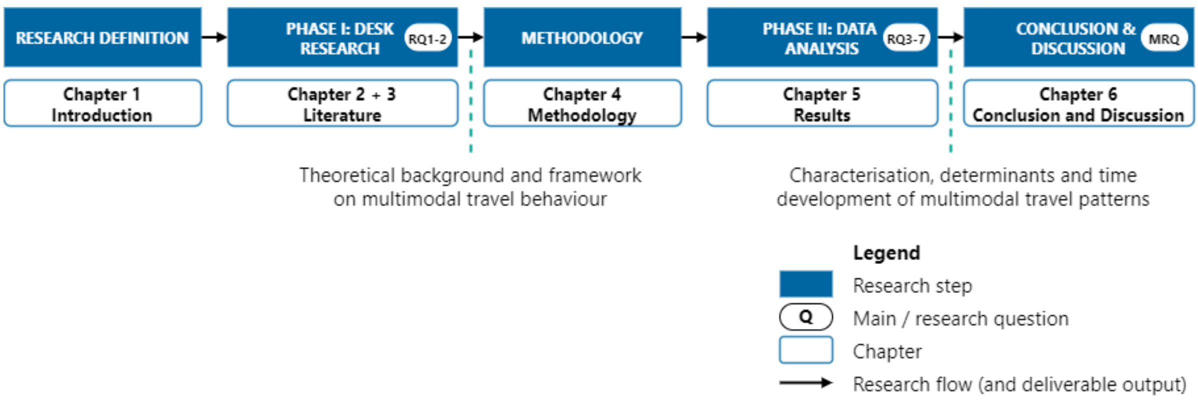


Figure 3. Research and chapter overview.

First, the research definition, including a synthesis of knowledge gaps on multimodal travel behaviour, is elaborated in this Chapter 1. This synthesis led to the research questions (Table 1), research overview (Figure 2), and research flow (Figure 2), with the general research approach and deliverables per research phase. Second and third, the approach for desk research (phase 1: RQ1-2) and the output of the theoretical framework, including background information about multimodal travel behaviour and determinants based on literature, are discussed respectively in Chapters 2 and 3. Fourth, before the main and second research phase (phase 2: RQ3-7; respectively sections 5.1-5.5) consisting of analysing the data with LCA, the methodology and viewpoint towards other methods are provided in Chapter 4. Moreover, that includes empirical data gathering, data operationalisation, processing of existing quantitative survey datasets, creating measurements and model estimation. Fifth, Chapter 5 presents the research outcomes and explains the defined travel user groups, determinants and time development of travel users' modal travel patterns in the results research step. Last and sixth, the research is concluded and discussed in Chapter 6, including a discussion regarding policy and scientific implications.

2. Literature study: overview and definition of multimodal travel behaviour measurements

This chapter aims to set the scope for the use of multimodality measures by overviewing, comparing and defining measurement categories and measures. The approach for analysing multimodal travel measurements (categories) is provided first (section 2.1). Second, the multimodality measurements are classified into categories (section 2.2), whereafter, thirdly, the (dis)advantages of the categories (in relation to the used method, section 4.1) are analysed (section 2.3). This is done to synthesise the findings to choose the scope of which categories, with corresponding measures, are suited for our research. Last, the multimodality measures within the chosen categories are analysed and synthesised (section 2.4) to define the measures used for this research.

2.1 Approach of classification and comparative literature study

As the general definition of (intrapersonal) multimodal travel behaviour (section 1.2) used in our study is set, several measurements of multimodality need to be explained. Our main study aim is to show the multimodal development of groups of individuals over time. For this, measurements, or so-called metrics or indices, are ideally suited to objectively differentiate among groups and acquire a comprehensive view of multimodality, as explained later (section 2.3/2.4). However, the guidelines on using which measures when are not yet agreed upon (section 1.2). Therefore, a comparative literature study is done of widely used multimodal travel behaviour measurements, which are positioned into the existing body of research, describing, comparing, reviewing, and using them. This section describes the approach of comparing the literature and how the literature is found.

A general classification of categories is set up for this (section 2.2). The suitability of categories and measures within the categories are assessed to set the scope and to choose and define measurements for this research (section 2.3/2.4). So, the perspective in this comparative analysis is to argumentatively compare the combination of differences, comparability and (dis)advantages between measurements leading to the measurement choice customised for our specific research purposes by integrating elements of the few existing (comparison) works. The analysis approach in this comparative analysis is thus not to thoroughly overview ('all') multimodal measurements and show in (mathematical) detail the differences between them but briefly position them.

The search approach is exploratory, as setting up a general framework (with room for extensions) is also more of an explanatory task[1]. However, conclusive information is, in the end, obtained about measurement choices. Because the classification is not the main aim of this research, and no refined general framework exists yet, an exploratory start is deemed sufficient. For the search approach, this means that a non-systematic approach suffices. Separate search terms (and synonyms), such as "travel behaviour", "multimodality", and "measurements" in several search combinations, are used. Literature is searched in several literature databases (Google Scholar, Web of Science, Scopus, ScienceDirect), in which appropriate peer-reviewed journal literature is selected, limiting the number of articles screened when having about ten valuable articles. Moreover, the approach consisted mainly of backward (and forward) snowballing to identify more relevant literature. Literature is appropriate when being in the existing body of research, as explained in the previous paragraph, and the measurements used are commonly used. In sections 2.2 and 2.3, the sources from this search are used when necessary, and the sub-approaches are explained. In section 2.4, it is explained how the previous ones identified (which are not generally overviewed) are potentially selected for comparison and what the sub-approach is.

[1] https://research-methodology.net/research-methodology/research-design/exploratory-research/

2.2 Classification: multimodality measurements in categories

Before diving into specific measurements, a brief classification framework is given of the different measurement categories of measures used in existing research to create our own classification.

Several distinctions and classifications exist in research, which are mainly the following: (1) predefined categorisations/conceptualisations, (2) data-driven classifications, and (3) continuous indices (e.g., An et al., 2021) or (1) nominal categories (1 and 2 of the previous) and (2) quantitative indicators (comparable to 3 of the previous) (e.g., Scheiner et al., 2016). Existing classifications have slightly different nuances. The measurements are put into a more refined and general framework (Figure 4). Still, other categorisations, for instance, with other groupings or more levels or categories (more exhaustive viewpoints), could suffice. Nevertheless, it might be a starting point for researchers when the aim is to create new taxonomies. The categories are mutually exclusive, as all measurements (used in the comparative analysis, sections 2.3/2.4) fit into one category without any overlap.

Broadly speaking, measurements can be divided into the following classification (Figure 4). Multimodality measurements consist of nominal characterisations (1) and numeric indicators (2). Predefined characterisations (1.1) and data-driven (1.2) classifications fall into the mentioned nominal characterisations category. Sub-categories of the predefined category are dichotomous characterisations (1.1.1) and polytomous characterisations (1.1.2). The numeric indicators can be divided into one-sided (count or continuous scales, 2.1) indicators and multi-sided (continuous scale, 2.2) indicators. The classification is subsequently described.

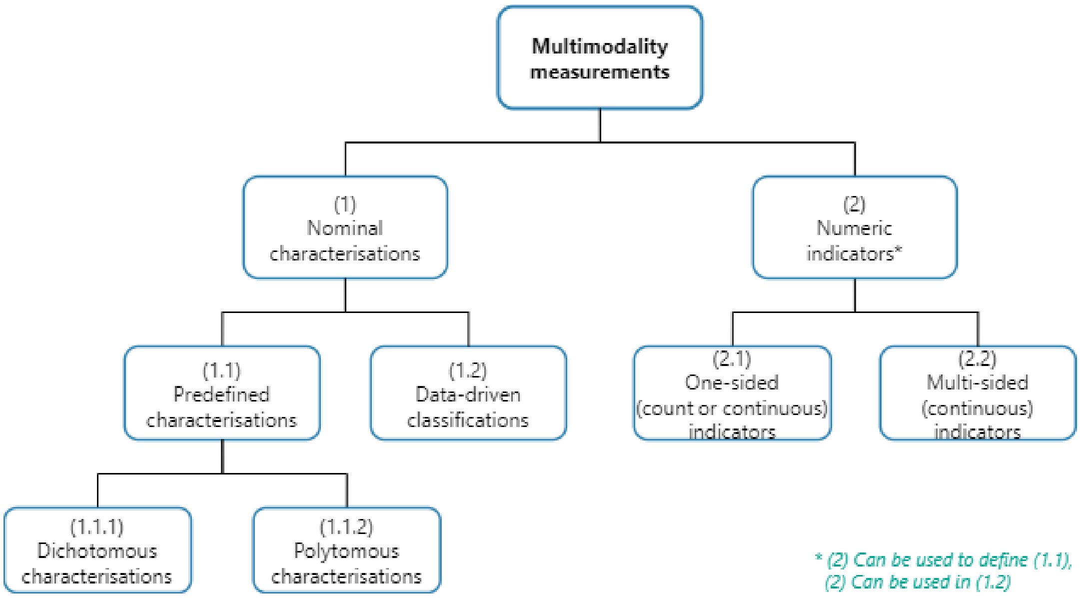


Figure 4. Classification and scope of multimodality measurements.

The multimodality measurements are all an attribute of individuals (or groupings) and their respective travel patterns consisting of trips and stages (part of trips) (e.g., Heinen & Mattioli, 2019a). For instance, a trip can comprise a cycling stage and a public transport stage. The categories are compared and explained according to Figure 4.

Nominal characterisations (1) are characterisations of combinations of used transport modes by individuals (Heinen & Chatterjee, 2015; Heinen & Mattioli, 2019a; Scheiner et al., 2016). Nominal characterisations mean that several nominal groups are identified. Other aspects than modal combinations can also be used to define groups (see sub-categories), which are sometimes modal intensities next to mode use only. Moreover, the numeric indicators (see category 2) can be used as a criterion for defining nominal groups. This category can be split down into (1.1) and (1.2).

Predefined characterisations (1.1): are conceptualisations which mainly look at the variety of combinations of modes used in a certain period (An et al., 2023; Heinen, 2018; Heinen & Mattioli, 2019a; Scheiner et al., 2016). Predefined means that beforehand, groups and criteria are defined to determine which individual is regarded to belong to which group. The notion of predefined should not be misunderstood with numeric indicators, for instance, where formulas are created beforehand. Groups might, for example, be in front characterised as 'monomodal car users', where individuals only used a car or 'multimodal car users', where individuals used a car and one or more modes of transportation, according to Heinen and Mattioli (2019a). So, being multimodal could depend on the mixture of using a sufficient amount of intensity of other modes (An et al., 2023). This category can be further divided into (1.1.1) and (1.1.2).

Dichotomous characterisations (1.1.1): are predefined characterisations which consist of two nominal non-overlapping groups. For instance, being multimodal (yes/no), based on certain criteria (Scheiner et al., 2016). The example given above is also dichotomous. Other two-sided categorisations found in the literature are, for instance, being multimodal when using more than one mode notwithstanding the frequency of use, being multimodal when no mode was used in more than 70% of trips (Nobis, 2007), or being multimodal when people used at least two out of three different transport mode groups (Kroesen, 2014).

Polytomous characterisations (1.1.2): are predefined characterisations which consist of at least three nominal non-overlapping groups. The operationalisation used as in the study of Heinen et al. (2018) is polytomous: unimodal car users (individuals only using the car), unimodal cyclists, uni/multimodal users (individuals only using the same mode, but not only the car or only the bicycle, or individuals only using the same combination of modes for every trip), and other multimodal users (individuals who varied their mode use between trips, or with a different multimodal pattern).

Data-driven classifications (1.2): are classifications determined and dependent on the data used (An et al., 2023; Heinen, 2018, Heinen & Mattioli, 2019a; Scheiner et al., 2016). This approach uses unsupervised classification methods (section 4.1); the nominal groups are thus emergent and not beforehand created based on criteria. The used clustering schemes are mainly k-means or latent class analysis to identify clusters with different travel patterns. The numeric indicators (2) can also be used as a characteristic to use in the clustering or to describe the clusters (see category 2). It incorporates mostly multi-sided characteristics, like the mode use and intensities of travel behaviour, compared to only using mode use by commonly used predefined characterisations. An example of emerged groups via clustering are the following named groups based on the identified clusters: car mostly, bicycle mostly, walk-car-bicycle, multimodal bus/tram/metro and multimodal train (An et al., 2022).

Numeric indicators (2) are quantitative measures to measure travel patterns' modal (intensity) use. These measures can be of scale count or continuous. Numeric indicators might also be used to base the predefined nominal groups on (1.1), and these indicators can be used together with data-driven classifications (1.2). The (2) category can be split down into (2.1) and (2.2).

One-sided (count or continuous) *indicators* (2.1): are quantitative measures which use one-sided characteristics of travel patterns. As seen in the predefined characterisations, this is most often also the case in that category, where mostly only modal use/combinations are considered without the modal intensities. An example of a one-sided count indicator is the number of modes used (Diana & Pirra, 2016). An example of a continuous one-sided measure is the share of the primary mode or the difference in the proportion of stages between primary and secondary modes (section 2.4).

Multi-sided (continuous) *indicators* (2.2), hereafter named 'continuous indicators' for being recognisable compared to the literature: are quantitative indicators, as opposed to nominal characterisations, which use the mode use variability with multi-sided characteristics for defining the level of multimodality (Heinen, 2018; Heinen & Mattioli, 2019a; Scheiner et al., 2016). These indicators consider the variety of modes used and intensity in several mathematical formulations, which can be seen as a generalised measure of modal habit (Diana & Pirra, 2016). A variety of indicators exist in which the variability is measured according to them. One example is the Herfindahl-Hirschman Index (balance in the distribution of mode usage), proposed by Heinen and Chatterjee (2015), see section 2.4.

2.3 Comparative study 1: multimodality measurement categories

As the measurement categories are classified (section 2.2), the (dis)advantages of these categories are now analysed (section 2.3.1), and the suitability of the measurement categories in relation to the chosen method LCA (section 4.1) is given (section 2.3.2). These findings are synthesised to decide which measurement category or categories are suited for this research (section 2.3.3).

2.3.1 Analysis: (dis)advantages of multimodality measurement categories

The measurement categories (section 2.2, Figure 4) are analysed based on their (dis)advantages. First, the nominal characterisations (1) are analysed. Then for the numeric indicators (2), firstly, the numeric multi-sided (continuous) indicators (2.2). Second, the numeric one-sided indicators (2.1) are analysed as the latter category exhibits some pros and cons comparable to the categories already mentioned. Last, using both numeric indicator sub-categories jointly is examined.

Nominal characterisations (1)

The nominal characterisations (1) can be interpreted intuitively as distinct groups can be identified (An et al., 2021). The definition of multimodality can be based on many criteria, including numeric indicators (2), and it is a straightforward method. Especially the predefined dichotomous characterisations (1.1.1) are intuitively and easily describing if persons are multimodal (or not).

Nevertheless, it is argued that nominal characterisations give no insight into the intragroup (i.e. between groups) differences and the levels of variability within groups, also named intrapersonal variability (An et al., 2023; section 1.2). In the predefined binary characterisations (1.1.1), being more or less multimodal is also not effectively captured, but data-driven classifications (1.2) can somewhat show where the clustering is based on. The difficulty in capturing the intragroup and within-group variability in the nominal category is because travellers are categorised into aggregate non-overlapping groups without capturing the individual level of multimodality (Heinen & Mattioli, 2019a). Moreover, predefined (1.1), data-driven (1.2) nominal groups are most closely related to the use of certain (combination) modes and not to the variability of modes (Heinen, 2018; Heinen & Mattioli, 2019a; Scheiner et al., 2016). So these characterisations might be more related to reporting if certain mode uses to increase or decrease instead of multimodality.

Numeric indicators (2)

Compared to the nominal characterisations (1), the numeric multi-sided (i.e. continuous) indicators (2.2) are thus more effective and likely to represent better the individual level (intrapersonal variability) of multimodality (Heinen, 2018; Heinen & Mattioli, 2019a), complementary to the modal variability only (e.g., Heinen & Chatterjee, 2015). The induvial travel pattern, modes variability, and intensity are better reflected in a balanced way among continuous indicators. Furthermore, several (quantitative) levels differentiate the extent of multimodality better than nominal groups (An et al., 2023). According to

Scheiner et al. (2016), this allows for inter-individual comparisons.

The downside is that the continuous indicators do not describe the use of specific modes and that it cannot be specified which modes should be used in order to be 'more' multimodal (Heinen, 2018; Heinen & Mattioli, 2019a; Scheiner et al., 2016). An example is given by Scheiner et al. (2016), where a 50-10-10-10-10-10% frequency distribution of modes used gives the same indicator value independent of which mode accounts for the 50% of the continuous indicators used. Moreover, interpretability is more of an issue with these measurements (Heinen, 2018).

The numeric one-sided indicators (2.1) have (dis)advantages from the (1) and (2.2) categories. The one-sided indicators are also regarded as intuitively and straightforward describing individuals, like the number of modes (An et al., 2023), just as the nominal groups. Moreover, the several (quantitative) levels can better differentiate the level of multimodality better, just as the continuous indicators.

However, these one-sided indicators are also closely related to using certain (combination) modes and not to the variability of modes that hampers interpretability. It provides only an overall indication of the extent of modal variability but takes no account of the frequency of use of different modes (Heinen, 2018; Heinen & Chatterjee, 2015).

Both numeric indicators (2) sub-categories

The joint use of both numeric indicator categories has advantages. Heinen (2018) explains that, based on the work of Diana and Pirra (2016), using multiple kinds of measures complements the research. Using both sides, namely intuitive (one-sided) and complex (continuous) indicators, aligns with the idea of using more complex measures for research purposes and more intuitive measurements for policy purposes (Scheiner et al., 2016), which can make these joint use of them suitable.

2.3.2 Analysis: suitability of multimodality measurement categories with LCA

After the categories are explained (section 2.3.1), the suitability of the method with them is explained. The method used in this research, Latent Class Analysis (LCA) (section 4.1), clusters individuals into emergent clusters based on the data attributes and has some measurement categories (section 2.2) which are suitable or inappropriate to be used (jointly). These are now explained.

When looking at the multimodality measurement classifications, clustering belongs to the nominal data-driven classification. So, the identified nominal groups of individuals with comparable patterns can measure the multimodality based on the level of the attributes in those clusters. Simultaneously, the nominal predefined characterisations are not able and thus suitable to be used in combination with LCA. Nevertheless, when applying LCA, numeric indicators can be formally used to describe better and understand the created clusters and their individuals. Both numeric one-sided and multi-sided (continuous) indicators are suitable for this.

2.3.3 Synthesis: positioning of multimodality measurement categories in this research

Based on the (dis)advantages of the measurement categories (section 2.3.1), together with their suitability with LCA (section 2.3.2), our study is positioned (Figure 5) in the multimodality measurement classification by synthesising the previously found.

As explained, when applying LCA (section 4.1), it belongs to the data-driven classifications. Nonetheless, the number of indicators can formally be added to describe the created clusters and their individuals. The combinational use of numeric indicators next to the data-driven classifications (nominal characterisations) can overcome the disadvantages of those (overarching) categories alone. Moreover, using both categories within the numeric category creates a higher value for research purposes and policy practices, making the numeric indicators suitable to use jointly.

The main disadvantage of not capturing the intrapersonal variability well of the individual travel pattern in nominal characterisations partly holds in the clustering approach. Since with LCA, travel behaviour indicators are chosen on which individuals are clustered, the different values on these indicators for different nominal groups can be shown. With this aggregate understanding, some individual information is lost. However, the advantage of the method is to intuitively identify underlying comparable groups in society, which can be understood via this method. The shortcoming of aggregation and capturing the induvial variability less accurately can partly be overcome when a numeric measurement, in which more individual characteristics are captured, is also used, although being grouped later. The disadvantage of the numeric measurement is that it is not able to show which exact travel modes are used, which is overcome by using LCA as the specific mode intensities can be shown.

When using measurements from both the numeric measurement categories, their limitations can be overcome separately. By definition, continuous indicators capture both the modal variability and the modal intensity and are deemed to measure the multimodality better; the disadvantage is the limited interpretability of those indicators. On the other hand, the one-sided indicators have the advantage of intuitive and straightforward describing individuals, like the nominal characterisations category.

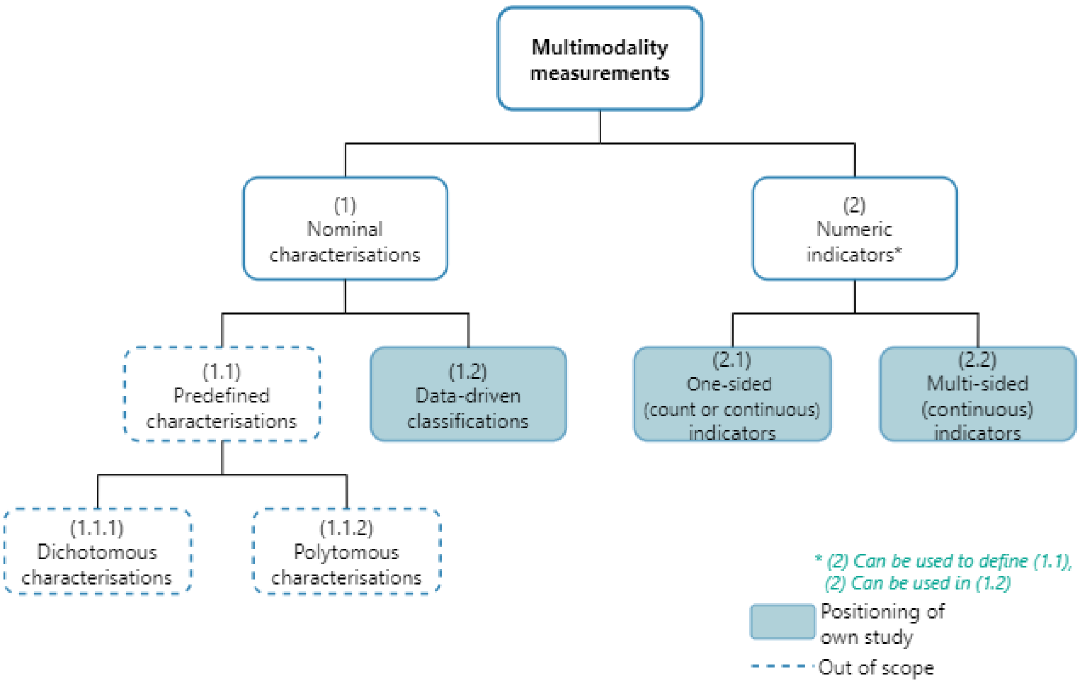


Figure 5. Classification, scope, and positioning of own study concerning multimodality measurements.

The positioning of this research is summarised (Figure 5). The research is about applying a data-driven method (1.2) to cluster people in nominal groups (1), where the focus is on using numeric indicators (2) from both categories (2.1 and 2.2). The predefined characterisations (1.1) are out of scope and not used to define multimodality.

2.4 Comparative study 2: multimodality measures within numeric indicators category

As explained in the synthesis of the multimodality measurement category analysis (section 2.3.3), both numeric indicator measurement categories will be used. This section is the second comparative study in this chapter to compare the measures within these categories, the one-sided and continuous indicators, to choose and define the multimodality measurements for this research. From each category, one measure will be selected to have a joint complementary view, overcoming potential shortcomings of both categories alone and remaining interpretable by not having too many indicators. This section first provides the literature selection and overview (section 2.4.1), whereafter, the analysis and synthesis are given in multiple sections. First, the analysis is given about whether the results of multimodality measures (for both numeric indicator categories) are comparable (section 2.4.2). The one-sided indicators are compared (section 2.4.3), and the continuous indicators are compared (section 2.4.4). Last, the synthesis is given (for both numeric indicator categories) to choose and define multimodality measures (section 2.4.5).

2.4.1 Literature selection and overview of numeric multimodality measures

The literature for comparison of numeric multimodality measures is selected and after the table described. The previously used and found literature is considered for this comparative study of multimodality measures. The source is included when at least two widely used measures (or variations based on them) are used. Moreover, the numeric indicators should be present. Additionally, it is not necessary if the source has compared the measures themselves. However, if more information from other sources is needed, these sources can be used in the following sections to give descriptions or argumentations on specific indicators. Nevertheless, these sources will not be used in this second comparative study itself.

Table 2. Overview of studies on multimodality measures with their characteristics.

Study Author	Context	Overall aim	**Comparative component***
An et al. (2022)	The Netherlands, 2018, MPN data	The distribution of mode-specific attitudes and attitude-mode use incompatibilities across clusters and levels of multimodality.	
An et al. (2023)	England, 2016, NTS data	The extent to which the level and correlates of multimodality vary by trip purpose.	✓
Diana & Pirra (2016)	Italy, 2012, ISTAT / simulation	Theoretical investigations and empirical/simulation experiments on the properties of multimodal indices.	✓ (review)
Heinen (2018)	The Netherlands, 2015, questionnaire	The extent to which multimodality is associated with changing mode usage in clusters.	
Heinen & Chatterjee (2015)	Great-Britain, 2010, NTS data	The extent to which people use mixtures of transport modes and the predictors of this.	✓
Heinen & Mattioli (2019a)	England, 1995-200, NTS data	The extent of multimodal trends.	✓
Scheiner et al. (2016)	Germany, 1994-2012, GMP data	The extent of changes in multimodality due to life course events.	✓

*The research identified similarities (or differences) between the used indicators.
Abbrev.: GMP = German Mobility Panel, ISTAT = Italian National Statistical Institute ('Aspects of daily life' survey), MPN = Mobility Panel Netherlands, NTS = National Travel Survey.

The selected literature is shown in Table 2, where the context and overall aim are mentioned. Some studies included a comparative component in the study next to the overall study aim. One study by Diana & Pirra (2016) is a review study that used an empirical dataset and simulation to compare measures. The identified literature is used in the upcoming sections.

2.4.2 Analysis: comparability of one-sided and continuous indicators

Table 3 shows the research which is using one or multiple indicators based on the literature selection (section 2.4.1).

One interesting observation is the study of Diana & Pirra (2016), which reinterpreted indices from literature based on welfare economics (a.o., social equity), information theory and ecology, to make it suitable for transport-related research on multimodality. They are regarded as one of the pioneers in systematically comparing the indicators based on representing the balanced mode use as well as the variability of it (see definition multimodality, section 1.2). In the table's last column, the synthesis of indicators' similarities shows that empirically, little changes are seen in the multimodality measuring of trends, clusters, or a population set. Moreover, the results of further analyses with the measures were hardly affected. This holds for the one-sided indicators, the continuous indicators, and between them. The stage-trip levels mentioned in the table are explained earlier and more in-depth in section 4.4.

Table 3. Overview of multimodality measures comparison.

Study Author	**Numeric indicators*, per category** (trip/stage level)	**Indicator comparison**
An et al. (2022)	Continuous (stage): • HHI • OM_PI	Clusters with the indicators have: o quite similar (opposite) values.
An et al. (2023)	One-sided (stage): • Number of modes • Difference between primary and secondary mode Continuous (stage): • HHI • OM_PI	Indicators among purposes have: o highly consistent results.
Diana & Pirra (2016)	Continuous (stage, most likely): • HH • GI • HH_m • ATK • OM_PI • DAL • OM_MI • DAL_m • TH	Indicators are: o all not mathematically outweighing the others on (un)desirable properties; o suboptimal for four relatively well performing on multicriteria assessment: • HHm • OM_PI • OM_MI • DAL_m
Heinen (2018)	One-sided: • Number of modes/modal combinations (trip) • Number of modes (stage) • Highest share of a mode/modal combination (trip) Continuous (stage): • HHI • OM_PI	Clusters and predefined groups with the indicators are overall speaking: o not contradictive; o quite similar in terms of proportions of high/low values.
Heinen & Chatterjee (2015)	One-sided (stage): • Number of modes • Number of stages • Difference between primary and secondary mode counts	Indicators show, overall speaking: o strong correlations with each other; o a moderate correlation to proportion of use of each mode category.

	Continuous (stage): • HHI 8 modes • HHI 3 modes	
Heinen & Mattioli (2019a)	One-sided (stage): • Number of stages • Difference between primary and secondary mode shares	Trends of all indicators are: o very similar.
	Continuous indicators (stage): • HHI • MM • HHm • DAL_m • OM_PI	
Scheiner et al. (2016)	One-sided indicators (trip, main mode): • Number of modes • Share of primary mode	Indicators have: o little contradiction.
	Continuous indicators (trip, main mode): • HHI • Shannon entropy	

*See formulas in literature and on what literature the formula is based/derived/reinterpreted. Same named indicators can thus be slightly differently operationalised.
Abbrev.: HH (Herfindahl-Hirschman) index, HH_m (modified HH) index, HHI (modified HH) index, OM_PI (Objective Mobility Personal Index; based on original Shannon entropy), OM_MI (Objective Mobility-level-sensitive index; based on OM_PI), TH (Theil entropy) index, GI (Gini) index, ATK (Atkinson) index, DAL (Dalton) index, DAL_m (modified DAL) index, MM (Multimodal) indicator.

2.4.3 Analysis: one-sided indicators

The extracted numeric indicators from the selected literature (Table 3, section 2.4.2) are shown in Table 4. For all measures, some are measured at the trip level and some at the stage level (stages are part of a trip, e.g., a trip can comprise a cycling and public transport stage). A measure at the trip level reflects the variability between trips, and, on the other hand, a measure at the stage level considers the variation among all used modes (Heinen, 2018). The measures (Table 4) are subsequently compared.

Table 4. Numeric indicators (multimodality measurement category).

NR	**Numeric indicators** One-sided	**Level** (trip/stage)	**Scale**	**Study*** Author(s)
1	Number of modes	Trip (modal combinations)	Count	Heinen (2018); Heinen & Mattioli (2019a)
		Stage	Count	An et al. (2023); Heinen (2018); Heinen & Chatterjee (2015); Heinen & Mattioli (2019a); Scheiner et al. (2016)
2	Number of stages	Stage	Count	Heinen & Chatterjee (2015)
3	Highest share of a mode/ modal combination	Trip	Continuous	Heinen (2018)
4	Share of primary mode	Trip	Continuous	Scheiner et al. (2016)
5	Difference between primary and secondary mode shares/counts	Stage	Continuous/ Count	An et al. (2023); Heinen & Mattioli (2019a); Heinen & Chatterjee (2015)

* One selected study, Diana & Pirra (2016), is not shown; it does not include one-sided indicators.
1-2: Higher value indicates a higher mode variability (higher 'multimodality').
3-5: Lower value/smaller difference indicates a lower dependence on the primary mode (higher 'multimodality').

Measures, like the number of modes, are more often used or suggested (NR1 Table 4). It captures the multiplicity of modes a traveller uses intuitively (An et al., 2023). Most literature used this measure on the stage level, and some on the trip level where the different modal combinations in a trip are used then. However, this measure is affected when measuring at the trip or stage level (e.g., Diana & Pirra, 2016). For the number of stages (NR 2), comparable to the number of modes (NR 1), it is likely to be assumed that someone with more stages (parts of trips) used more modes. For instance, if someone uses a car for a trip, one stage is sufficient to reach the destination. However, it is no direct measure of modal variability. The latter measures (NRs 3-5) are mainly about shares of mode usage or differences in use between primary and secondary mode shares/counts (on the stage level). It can effectively capture the dependence on a mode (An et al., 2023; Scheiner et al., 2016), and it is argued that a multimodal, more balanced user has a lower dominant mode use. So, these measures provide less information about variability among the full travel pattern or timespan, as only the primary and or secondary mode is considered.

2.4.4 Analysis: continuous indicators

In contrast to the numeric indicators, where all identified ones (section 2.4.2) are compared (section 2.4.3), the continuous indicators are more diverse and measure all something different. So, a limited amount of indicators is further compared to not redoing already done comparative work and making sure more commonly used measures are considered for travel behavioural research.

The most used continuous indicators (Table 3) are OM_PI and HHI, which are both also recommended by Diana & Pirra (2016). Both are commonly used in capturing multimodality for travel behavioural research, as Heinen & Mattioli (2019a) explained for HHI and An et al. (2021) for OM_PI. Some comparisons are given for them.

Both measures have the desirable property of being replication variants (Diana & Pirra, 2016). This means that when someone cycles ten times a day and someone else cycles and drives both ten times a week, the latter will be calculated as having higher multimodality (An et al., 2022). So the multimodality index will not remain the same when replicating given modes with their intensities (An et al., 2023). Moreover, due to replication variance, these measures are suited to measure multimodality by considering the number of modes from the predefined set, irrespective of whether these modes are available (Diana & Pirra, 2016). They also highlight that this problem of non-availability is also visible in the choice set definition of mode choice models and is easily encountered in this way. It should be kept in mind that some seem less multimodal because they have access to a smaller number of travel modes. Nevertheless, when considering the number of travel means that are used from a broader set, the 'real' multimodality is measured.

The HHI, used by Susilo and Axhausen (2014), is especially useful when examining the repetitiveness of travel behaviour combinations over a period. This is because it is derived from market concentration, where more weight is put on bigger markets and, in our case, the repetition of travel patterns. This means that persons who have a higher concentration of stages among some modes are less multimodal (Heinen & Mattioli, 2019a). On the other hand, lower specific concentrations (and a more balanced distribution) mark a more flexible or less repetitive traveller, which is considered more multimodal (Heinen & Chatterjee, 2015).

The OM_PI is based on the Shannon entropy formula, as explained by Diana & Pirra, 2016. It is grounded in information theory, and this concept is reinterpreted for travel research purposes into OM_PI by Diana and Mokhtarian (2008, 2009). OM_PI measures variability. A higher value means a higher level of variability in mode use and greater multimodality (Heinen & Mattioli, 2019a). Scheiner et al. (2016) explain that this measure describes the amount of heterogeneity in the distribution of modes.

Moreover, they explained that entropy (where OM_PI is based on) puts more weight on weakly considered modes compared to HHI. An example given by Diana and Pirra (2016) shows two travellers, A and B, which have intensities of mode use [10 10 3 0] and [14 8 1 1], in which the index value takes 0.29 for A and 0.32 for B. The latter traveller, B, has a slightly higher multimodality because of a higher number of modes (more diversity). But individual A is almost equal multimodal with less modal variability but somewhat more equal divided intensities (equality).

2.4.5 Synthesis: numeric indicators choice

Based on the previous analysis for both numeric indicator categories (sections 2.4.3 and 2.4.4), the following measurements are chosen based on argumentations in this synthesis and afterwards explained.

For both measurement categories, hardly any differences are visible in the results of other studies which use measures of these categories (section 2.4.2). For the one-sided indicators, the number of modes is chosen. This one is used because the number of modes is widely used and captures the actual travel patterns (instead of looking at shares of primary modes, e.g.). Moreover, this measure is more intuitive. Concerning the continuous indicators, both promising continuous indicators are widely used (HHI and OM_PI). The choice to use one measure can be based on which measure is more suited for our particular research. The HHI is more suited for the repetitiveness of travel patterns, while OM_PI can effectively capture diversity (a variety in mode usage) and equality (a balanced mode use with equal frequencies). Although both measures showed similar results when measuring in empirical research, our research objective of identifying the variability of modes is more in line with OM_PI. This seems more suitable as it better captures the fullest variability (also of weakly considered nodes, as explained in section 2.4.4). This is important for our scope of analysing the variety in one specific timeframe instead of a measure which puts more weight on the repetitiveness.

The chosen numeric indicators are summarised:

Multimodality measurements choice

One-sided:	Number of modes
Continuous:	Objective Mobility Personal Index (OM_PI)

The number of modes is intuitively measured, and OM_PI is explained in the paragraph below.

OM_PI ranges from 0 to 1, respectively, the exclusive use of one mode and where all modes are equally used in the same intensity (An et al., 2021; not part of this comparative analysis). So, a higher value means a higher degree of multimodality. As explained by Diana and Pirra (2016), it can be used for the amount of trip(s) (segments), distances and travel times, so they propose the general notion intensity of use (f). The choice of level of intensity, as well as the level of analysis, is given in section 4.4. The mathematical formulation is based on this notion of intensity of use and on the formula provided by An et al. (2022), see (Equation 1). OM_PI is calculated for each individual i. It takes the share of the intensity of use (f) by a specific mode (j) and the modes of the whole set (N) into consideration. The equation is not further explored in mathematical depth, but the characteristics of measuring multimodality are already explained. Moreover, what the measured values of OM_PI means are further explained in the results (Chapter 5, section 5.2).

OM_PI_i =

$$\sum_{j=1}^{N}\left(f_{ij} * \ln\left(\frac{1}{fij}\right) * \left(\frac{1}{\ln N}\right)\right)$$

(Equation 1)

OM_PI_i Objective Mobility Personal Index for individual i

f_{ij} Share of intensity of use by specific mode j for individual i

N Amount of modes in sample considered

3. Literature review: overview of multimodal travel behaviour determinants

This chapter is about performing a literature review on multimodal travel behaviour determinants. In order to do so, the literature review approach is given first (section 3.1), including the search strategy. Second, the literature is overviewed and compared, and an explanation is given of how the synthesis is structured (section 3.2). Finally, the synthesis of results from studies about (multimodal) travel patterns is provided (section 3.3).

3.1 Approach of literature review

A systematic literature search process (provided in this section) identifies and synthesises potential determinants of multimodal behaviour. As various (not extensive) research is done about determinants, this systematic approach is suited to acquire a systematic and comprehensive view of determinants. A literature review is helpful when few extensive reviews have been performed yet (Van Wee & Banister, 2016). The search process is subsequently explained in this paragraph.

The systematic literature search process to capture determinants of multimodality consists of acquiring the first articles in Scopus using a search term. Scopus is a well-known and widely used literature database, including many top-ranked (transport-related) journals. This database is suited to find general literature about multimodal(ity) and travel (behaviour) research. The final search string, including Boolean operators, is set up after experimenting with separate search queries in Scholar Google and looking for potentially relevant articles. The search string is used in the TITLE-ABS-KEY field in Scopus to capture journals which include the search elements in the leading and most important fields. The search was conducted in May 2023, with the final search string, which accounts for several notations:

multimodal* AND travel.

Based on top-15 rankings (e.g., Journal Citation Report by literature database Web of Science), often listed journals from articles from the experimental search, and appropriateness of the journal contents to our research field, specific peer-reviewed journals are used to limit the number of search results. For example, journals which involve mainly travel behaviour research are included, whereas journals, e.g., primarily focussing on mathematics or solely about public transportation, are excluded. The following journals are used to limit the initial results in Scopus:

- Transportation
- Transportation Research Part A: Policy and Practice
- Transportation Research Part F: Traffic Psychology and Behaviour
- Transport Policy
- Transport Reviews
- Journal of Transport Geography

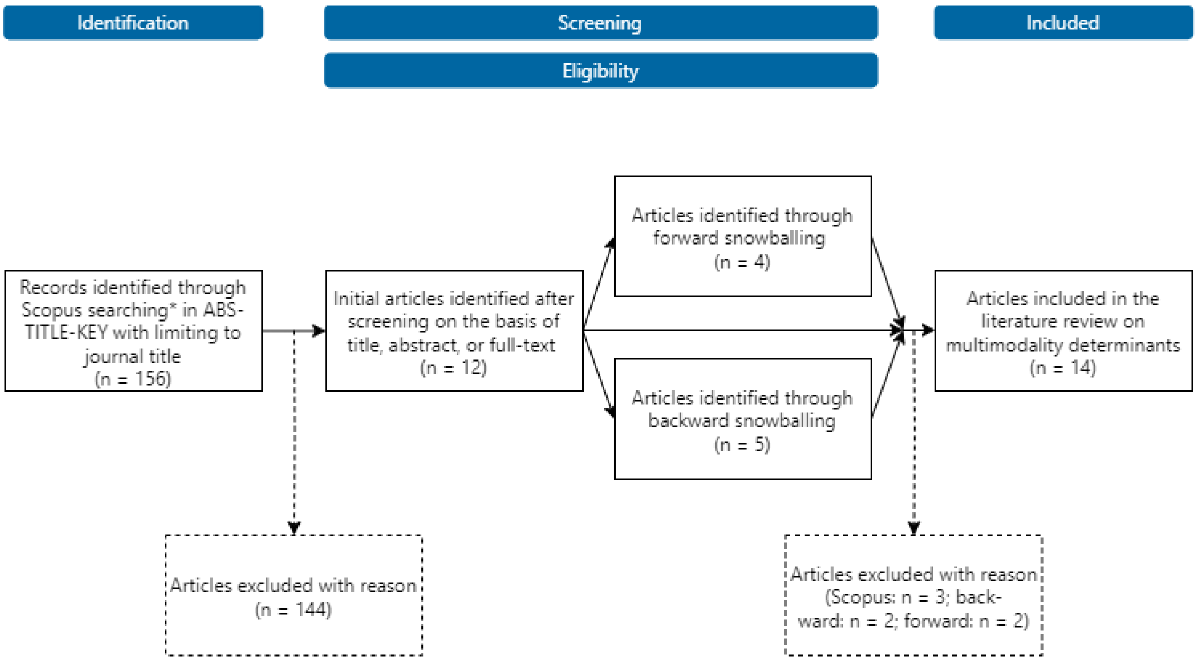
The search term and journal refinement resulted in 156 identified records. The search process is summarised in Figure 6, together with the final search string and the upcoming elaborated search and selection process.

The identified articles were screened by title and potentially with additional criteria content-wise in the abstract or full-text if they likely focussed on travel behaviour (of several modes) of individuals and the determinants of it. This implies that articles are eligible when covering individual travel behavioural patterns of several modes and assessing im/explicitly the multimodality (measuring) and determinants. Articles were also included if they covered the aspects of travel behaviour and determinants as part of their research. For instance, if the primary research aim is different (e.g., investigating attitude changes of multimodal travel segments). However, articles about, for instance, a change in travel behaviour due to Covid-19 or particular policies' effectiveness in measuring travel behaviour are not selected. Other criteria are also considered.

First, travel behaviour should be examined in a 'broader' sense, not only looking at leisure travel to give an example. This way, the regular day-to-day patterns are included, generally providing better information about multimodality. Some studies might be splitting results up for several travel purposes, which is also adhered for including. Moreover, articles studying the population in a 'broader' sense are included. So, e.g., articles focussing on predefined social milieus, employers or millennials only are excluded.

Second, the analysis of travel behaviour should be mainly based on the travel behaviour itself (like by travel times, distances or the number of trips with specific modes) to examine the actual usage instead of the likelihood or preferences (to change) of the way of travelling, for instance. Nevertheless, when travellers are grouped in research, the travel behaviour patterns can be examined together with attitudes but not with household characteristics for revealing household styles. So some studies are excluded, as it is expected that certain variables can be used to describe travel patterns without them being part of the patterns themselves, as the notion of modality styles differs when many household characteristics are already included next to travel patterns.

Third, about the travel patterns, it should, at minimum, include three (grouped) modes because bimodality is not always fully seen as multimodality. Moreover, it should include mode variability as opposed to travel time variability, e.g., (section 1.2), and being about aggregated trips as opposed to only considering trip chain complexities to match the definition as explained in section 1.2.



*Final search string:
TITLE-ABS-KEY (multimodal* AND travel) AND (LIMIT-TO (EXACTSRCTITLE , "Transportation") OR LIMIT-TO (EXACTSRCTITLE , "Transportation Research Part A Policy And Practice") OR LIMIT-TO (EXACTSRCTITLE , "Transportation Research Part F Traffic Psychology And Behaviour") OR LIMIT-TO (EXACTSRCTITLE , "Transport Policy") OR LIMIT-TO (EXACTSRCTITLE , "Transport Reviews") OR LIMIT-TO (EXACTSRCTITLE , "Journal Of Transport Geography"))

Figure 6. Summary of the article search process for reviewing multimodality determinants.

These explained criteria resulted in 12 initially selected articles after the screening, as seen in Figure 6. The initial articles were further screened via backward (in-text) and forward snowballing (first two pages on Google Scholar) to capture articles that might be missed due to the structure of the search terms used. Based on the explained additional criteria, five older articles cited by the initial articles and four newer articles citing the initial articles were selected. In total, 21 articles were initially eligible. However, seven articles are excluded from this set to limit the number of articles further. The articles are excluded when the analysis is based on predefined characterisations of the multimodal travel behaviour patterns of individuals. Although the articles are still feasible to examine determinants of travel patterns, the articles do not match the chosen scope in section 2.2. They are thus less comparable for our research. After all, 14 articles are included in the literature review on multimodality determinants.

3.2 Overview and structure of literature on multimodality determinants

The literature is overviewed (and compared) first, whereafter the structure for the review (section 3.3) is given.

Overview of literature

Based on the literature search process in section 3.1, the resulting studies are overviewed with characteristics in Table 5. In addition, the studies are grouped per multimodality measurement category according to the category definitions in section 2.2. Furthermore, the (main) method used and the study information (author, year and country) are shown. Finally, in the Country Code (CC) column, additional information is given via asterixis, whether nationally administered surveys, existing online panels or (smaller) questionnaires are used to show the extent of the study in general.

Table 5. Overview of reviewed studies on multimodality determinants with their characteristics.

Multimodality		**Study**			**Method**[c]
measurement category		Author	Year	CC[a,b]	
Numeric indicators	One-sided and Continuous	Heinen & Chatterjee	2015	GB	Multivariate regression
		Heinen & Mattioli	2019a	GB	Multivariate regression
		Scheiner et al.	2016	DE	Multivariate (change) regression
	Continuous	An et al.	2023	GB	Multivariate regression
		Susilo & Axhausen	2014	DE	Multivariate linear regression
Nominal characterisations	Data-driven: without Numeric indicators	De Haas et al.	2018	NL	Latent class (transition) analysis
		Kroesen	2014	NL	Latent class (transition) analysis
		Molin et al.	2016	NL[1]	Latent class analysis
		Olafsson et al.	2016	DK[2]	Cluster analysis
		Schneider et al.	2021	NL	Latent class analysis
		Ton et al.	2020	NL	Latent class analysis
	Data-driven: with Numeric indicators	An et al.	2022	NL	Cluster analysis
		Diana & Mokhtarian	2009	US[2]	Cluster analysis
		Haustein & Kroesen	2022	DK[1]	Latent class (transition) analysis

a: Country Codes (CC): alpha-2 ISO 1366.
b: National administered surveys are used unless other specified.
1: Existing online panels are used. / 2: Questionnaires (smaller surveys) are used.
c: Not necessarily the main method, only the one suited for our review is presented.

When looking at the studies to be used for reviewing, in Table 5, the differences between studies are first acknowledged as this makes one aware of the comparability matters before synthesising potential determinants. The studies are compared by the most prevalent differences, the multimodality measurement category used and the used context.

When looking at the multimodality measurement categories, a point of care is that there are differences between studies. Some studies involved nominal (multimodal) travel pattern groupings of individuals, with or without a multimodal measurement (with methods LCA or clustering, see section 4.1). In contrast, others analysed the influence of a numeric multimodal indicator (via regression analysis). Furthermore, other measurements, including other (grouped) travel modes, can be used within these categories. These differences make the results less comparable. For instance, when comparing studies, one named group, 'multimodal PT-based individuals', might not be another with the same name based on clustering, e.g., in terms of modal intensities of certain modes or model assumptions. Still, the general compatibility provides an intuitive way of exploring potential effects.

Besides the mentioned differences, studies also involved different contexts and study designs, i.e. smaller or bigger parts of countries, a smaller or bigger amount of representative participants, and different timespans. Regarding context, it can be seen that all studies (in this set) involve European countries or the US, which are all Western-oriented, meaning they are likely comparable in terms of wealth and Western cultures. Still, it is argued that people in the US are more car-dependent. In contrast, active travel and PT use are more prevalent in other Western countries, especially in cycling-oriented countries (e.g., Buehler, 2011). By using a literature set with several contexts, it is assumed that most variables to be analysed are chosen based on several studies describing the effect to limit the choice of variables based on only uniquely contextual effects of variables.

Structuring the literature review

In order to structure the synthesis on multimodality determinants, several overarching variable groups and accompanying variable groups are made. The following groupings are based on own interpretation, found in the literature studying multimodal travel behaviour (e.g., Molin et al., 2016), and based on the groupings of An et al. (2023), which drew upon the work of Heinen and Chatterjee (2015) with constraints of several domains on intrapersonal modal variability and Hägerstand's work in 1970 about constraints of spatial travel behaviour. It comes down to the following overarching groups: socio-demographic variables (including household-related variables), mobility resource variables and built-environment variables, from which accompanying variable groups are found in the literature (Table 6).

Based on the variable groups, it can comprise one or more variables and multiple operationalisations regarding measuring or transforming the same phenomenon can be used. Still, the results of all studies involving different operationalised variables or countries, et cetera, can be somehow generalised to explore insights about potential determinants of multimodality in our specific context, with the note that those studies are not comparable one on one exactly as earlier explained.

Table 6. Overview of overarching variable groups and variable groups of multimodality determinants.

Overarching variable group	Variable groups
Socio-demographic variables	Personal characteristics
	Household status
	Employment status
	Economic status
Mobility resource variables	Mobility resources
Built-environment variables	Built-environment characteristics

3.3 Literature synthesis on multimodality determinants

The studies are synthesised and marked in blue in Table 7 based on whether certain variables within variable groups (Table 6) showed associations with multimodality. When significance is given in the studies, only the significant variables, based on their significance criteria used in that study, are incorporated and synthesised in the upcoming sections. They are also grouped based on the multimodality measurement category. When reading the synthesis is important to know that a higher or lower multimodality measurement means that the level of multimodality was higher or lower, irrespective of how multimodality is measured in that study. Based on the overarching variable groups, the results on associations of variables with multimodal travel behaviour are synthesised per variable group (sections 3.3.1, 3.3.2, 3.3.3). The final set of variables (within the variable groups) and how they are operationalised is given in Chapter 4.

3.3.1 Socio-demographic variables

Personal characteristics, household status, employment status and economic status belong to the overarching socio-demographic variable group. These results about (multimodal) travel patterns are subsequently explained.

Table 7. Overview reviewed studies on multimodality determinants and variable groups.

Multimodality measurement category		Study Overarching variable group:	Variable group* Socio-demographic				Idem	Idem
			Personal characteristics	Household status	Employment status	Economic status	Mobility resources	Built-environment characteristics
Numeric indicators	One-sided and Continuous	Heinen & Chatterjee (2015)						
		Heinen & Mattioli (2019a)						
		Scheiner et al. (2016)						
	Continuous	An et al. (2023)						
		Susilo & Axhausen (2014)						
Nominal characterisations	Data-driven: without Numeric indicators	De Haas et al. (2018)						
		Kroesen (2014)						
		Molin et al. (2016)						
		Olafsson et al. (2016)						
		Schneider et al. (2021)						
		Ton et al. (2020)						
	Data-driven: with Numeric indicators	An et al. (2022)						
		Diana & Mokhtarian (2009)						
		Haustein & Kroesen (2022)						

*Only the significant variables within the variable group (when the method allows that) are included.

Personal characteristics

Personal characteristics comprise gender, age, ethnicity and education level, which are now elaborated per variable to find the relation with multimodal travel behaviour.

Gender

The association of gender with multimodality is explained for females first, then for males, and last, for both, together with a short conclusion.

Females are likely characterised as light (not multimodal) travellers when being unemployed or working in the household (Kroesen, 2014; Olafsson et al., 2016). Moreover, most bicycle classes have the highest shares of females (An et al., 2022; De Haas et al., 2018; Haustein & Kroesen, 2022; Schneider et al., 2021). However, in other studies, females (when working) have a higher multimodality measurement in general or for work trips only (An et al., 2023; Heinen & Chatterjee, 2015; Heinen & Mattioli, 2019a), or they belong more likely to a multimodal car-walk-bicycle class (An et al., 2022; Schneider et al., 2021; Ton et al., 2020), a multimodal bicycle class (Molin et al., 2016) or PT-based class (An et al., 2022).

Quite similarly, different findings exist for males, leaving room open for further research. Many studies show that comparable strict car users groups across studies have a higher share of men (An et al., 2022; De Haas et al., 2018; Haustein & Kroesen, 2022; Kroesen, 2014; Molin et al., 2016; Olafsson et al., 2016; Ton et al., 2020). On the other hand, males are also (somewhat) more visible in multimodal car-dominated groups (Diana & Mokhtarian, 2009; Molin et al., 2016) or more represented in PT-based multimodal groups (Haustein & Kroesen, 2022; Molin et al., 2016).

As opposed to the previous finding about PT and males, another study showed an equal gender split in a multimodal PT-based group (Olafsson et al., 2016). In addition, exclusive cyclist groups are also more balanced in some studies (Haustein & Kroesen, 2022; Olafsson et al., 2016). These and the previous contradicting findings show that more knowledge is needed concerning gender and multimodality.

Age

The association of age with having multimodal travel patterns is explained for older age groups, middle-aged groups and younger age groups. Ultimately, the findings are concluded for the age variable in general.

In many studies, elder age groups (above 60 or 65) showed lower multimodality measurements (An et al., 2023; Heinen & Mattioli, 2019a; Susilo & Axhausen, 2014). For example, being 60+ showed a significant decrease in multimodality measurement compared to a younger age group (Heinen & Chatterjee, 2015). In the study of Olafsson et al. (2016), elder age groups are more likely to be in a not-so-multimodal limited transport class. In another study, the effects of older age groups were less clear (Scheiner et al., 2016). Contradicting results to the previous can be seen in other studies as well, where multimodal (car-bicycle or car-walk-bicycle) groups were characterised by many retired people (An et al., 2022; Schneider et al., 2021; Ton et al., 2020).

Concerning middle-aged groups, those people are mainly associated with unimodal car-based groups, which are often quite large (An et al., 2022; Haustein & Kroesen, 2022; Olafsson et al., 2016; Schneider et al., 2021; Ton et al., 2020).

On the opposite, younger age groups are more multimodal measured compared to older age groups (Heinen & Mattioli, 2019a) and are associated with multimodal public transportation-based groups (An et al., 2022; De Haas et al., 2018; Haustein & Kroesen, 2022; Schneider et al., 2021; Ton et al., 2020). However, more unimodal bicycle-based groups or diehard cyclists are sometimes also characterised by younger adult ages (Haustein & Kroesen, 2022; Kroesen, 2014; Olafsson et al., 2016).

The previous shows that the effect of being older or younger on having a multimodal pattern is not always apparent. Moreover, the effect of middle-aged groups is mainly analysed in studies which involve only a measurement (which gives no insight into the actual travel pattern) or in clustering studies (which gives insight into the actual travel patterns but no insight into the multimodality measurement.

Ethnicity

Concerning ethnicity and multimodality, two studies can be analysed, whereafter, it is concluded.

A study by Heinen & Mattioli (2019a) showed that the differences in multimodality measurement increased between ethnic groups over time. Quite similarly to that study, the level of multimodality of the 'other ethnic groups' (compared to the white ethnicity group) had a slightly negative impact on the level of multimodality in the study of Heinen & Chatterjee (2015).

The preceding shows that only limited information about this effect is available and only operationalisations about being 'white' or not being used. This is often not the operationalisation used in Dutch (our scope, section 1.7.2) collected data, where information is gathered about whether being an immigrant. It contributes to a lack of knowledge about the immigrants' origin category in relation to multimodal travel behaviour.

Education level

The association of education with multimodality is explained in the following: first, associations with low multimodality, then for high multimodality and last, a conclusion is given.

Several classes consisting of light- or more heavy travellers with mainly car use (Molin et al., 2016), are characterised by a lower education level, shown in studies by De Haas et al. (2018), Diana & Mokhtarian (2009) and Kroesen (2014). However, educated people were in these studies also more likely to be in a car-dominated class. Moreover, students with yet low education levels are often seen in strict bicycle user classes (Kroesen, 2014; Schneider et al., 2021; Ton et al., 2020).

However, in the two cases already mentioned (Diana & Mokhtarian, 2009; Kroesen, 2014), educated people were also in a class with the highest multimodality measurement, meaning that the car is complemented with other modes but dominated by the car (De Haas et al., 2018; Diana & Mokhtarian, 2009; Haustein & Kroesen, 2022). Respectively, a more substantial increase in multimodality measurements was found for higher educated people in the study of Scheiner et al. (2016). Moreover, multimodal PT-based classes are also characterised by more educated people, and often students (De Haas et al., 2018; Molin et al., 2016; Schneider et al., 2021; Ton et al., 2020).

As shown, educated and not-yet highly educated (students) individuals are likely to be in less but also in more multimodal classes. The not yet crystalised effect of those education categories on multimodal behaviour makes it interesting to analyse them.

Household status

Household status and the effect of having multimodal travel patterns are now analysed, and a conclusion is given afterwards.

In previous studies, different findings exist regarding household composition in terms of the number of household members. Concerning the number of household members, having more is associated with belonging more likely to a less multimodal group. So, there is a higher likelihood of being in a more strict car user group as an individual in a bigger household (De Haas et al., 2018; Schneider et al., 2021; Ton et al., 2020). In contrast, a 2-person household is likelier to be in a multimodal 'car-walk-bicycle' class, as shown by Schneider et al. (2021) and Ton et al. (2020). The opposite holds for 1-person households, where it is more likely to be in more multimodal-based groups. On the other hand, over a third of respondents in the not-so-multimodal 'mainly bicycle, with occasional use of car' class is part of a 1-person household, according to De Haas et al. (2018).

Previous studies have thus shown a (in general) clear relationship between having more household members and less multimodal behaviour. The association between multimodality and 1- or 2-person households is less clear. Moreover, the mainly used operationalisation of the number of household members does not give a broad view of household status, as the kind of household composition would do. It makes it suitable for analysing household status from another perspective.

Employment status

Regarding employment status, several individuals' occupations are associated with (non-) multimodality. Studies involving measurements and/or data-driven classifications showed similar results. However, some unique findings exist for several statuses, which are now synthesised for non-working or non-full-time working individuals, students and being employed to give a conclusion on this variable.

Regarding individuals not working full-time, several studies show contrary findings. For example, working part-time (An et al., 2023; Heinen & Chatterjee, 2015; Heinen & Mattioli, 2019a), being unemployed (Heinen & Chatterjee, 2015; Schneider et al., 2021), or being retired (Heinen & Chatterjee, 2015; Heinen & Mattioli, 2019a) compared to fulltime working is associated with a higher multimodal measurement. Furthermore, in line with the previous, some studies find unemployed and/or retired people in a class with multiple modes or even diverse mode usage (De Haas et al., 2018; Schneider et al., 2021; Ton et al., 2020), and one study found the highest multimodal measurement for those people in a particular multimodal class compared to other statuses (An et al., 2022). However, Olafsson et al. (2016) explained that many retired and unemployed people are in a light traveller group with mainly unimodal behaviour. Also, individuals working in the household (mainly females), part of a subgroup of unemployed people, are likelier to belong to a so-called light traveller group, which also involves little travel behaviour (Kroesen, 2014).

Students, also not working full-time, showed overall seen multimodal behaviour. For instance, students in a traineeship or education level are also more multimodal measured, according to Scheiner et al. (2016). Multimodal behaviour among students is also visible, especially in classes with a combination of PT and other modes (An et al., 2022; De Haas et al., 2018; Kroesen, 2014; Olafsson et al., 2016; Ton et al., 2020), whereas students are also more likely to belong in somewhat less multimodal groups which mainly uses the bicycle (Kroesen, 2014; Ton et al., 2020).

On the contrary, being a worker negatively affected the multimodality measurement involving both activities and mode variability in a study by Susilo & Axhausen (2014). Working full-time or having a paid job is associated which classes which are not multimodal and are mainly car-based (An et al., 2022; De Haas et al., 2018; Kroesen, 2014; Molin et al., 2016; Olafsson et al., 2016; Schneider et al., 2021; Ton et al., 2020). On the other hand, some studies also find some higher shares of working people in more multimodal PT-based groups (Olafsson et al., 2016; Ton et al., 2020).

Many studies involved the employment status of individuals in studies analysing multimodality. It shows that many occupational categories are visible in both multimodal and non-multimodal classes. The specific effect of these variables on being in those classes remains to be determined.

Economic status

Income has several associations with being multimodal. First, the associations of lower incomes are explained, and second, the associations with higher incomes, to provide a synthesis of them both in the end.

Regarding lower income quintiles, compared to being in the highest income quintile, individuals have lower levels of multimodality (Diana & Mokhtarian, 2009; Heinen & Mattioli, 2019a). Moreover, individuals in the lower income quintiles reduced their multimodality more over several years, according to the study by Heinen and Mattioli (2019a). In addition, some found that unemployed individuals with lower incomes are more associated with less multimodal or unimodal classes (Kroesen, 2014; Olafsson et al., 2016). Nevertheless, individuals with lower incomes can also be associated with being in a somewhat multimodal class with car or bicycle use complemented with other modes (An et al., 2022; Kroesen, 2014; Molin et al., 2016). Students or others with, on average lower incomes are conversely more associated with multimodal behaviour classes, where PT is the primary mode of transportation (De Haas et al., 2018; Kroesen, 2014; Molin et al., 2016; Olafsson et al., 2016). These individuals can, however, also be associated with (more unimodal) strict bicycle user classes (Kroesen, 2014).

Regarding higher incomes, studies by An et al. (2022) and Diana & Mokhtarian (2009) found a class with the highest level of multimodality, characterised by multimodal behaviour, including PT, where individuals with higher incomes on average are more likely to be in this class. Compared to the lowest category, being in a higher or the highest household income category positively affects the multimodality measurement in another study by An et al. (2023) and Heinen and Chatterjee (2015). However, belonging to a higher socioeconomic status had no significant relationship with modal variability in the latter study. Besides higher incomes being associated with higher multimodality, high-income individuals are also more likely to be associated with classes characterised by strict car users, which are, in most cases, quite big groups (An et al., 2022; Kroesen, 2014; Molin et al., 2016; Olafsson et al., 2016).

Overall, individuals with higher incomes are more likely to be in more multimodal classes than those with lower incomes, with some exceptions, which makes it interesting to use in our specific research context.

3.3.2 Mobility resource variables

Mobility resources are mainly investigated in the literature on ownership of car licensures, cars, and bicycles. The association of mobility resources and mobility is explained for car ownership (including licensure) and bicycle ownership.

It should be noted that much research highlights that mobility resources are likely endogenous to travel patterns. This implies that, due to self-selection, someone that likes to have a specific mobility style is likely to acquire the several modes to be used (the other way around than travel models generally assume). So, owning several mobility options cannot be regarded entirely as exogenous in determining the specific model patterns (see section 4.2). By explaining the effects of those variables, it should be kept in mind that some research included these variables directly in the model, and others used them only as descriptive variables to avoid endogeneity issues, as no direct effects can thus be derived.

Car (licensure) ownership

Multimodality and the relation with 'cars' are also investigated in the literature. The relationships with multimodality are elaborated for having a car driving license and car ownership, and they are finally concluded.

Holding a driving license is associated with lower multimodality measurements (An et al., 2023; Heinen & Chatterjee, 2015). However, Scheiner et al. (2016) found an increase in the level of multimodality only for females owning a driving licence. Overall speaking, it is also found that when having a driving licence, one is more likely to be in a monomodal, less active traveller, or car-based group, and on the opposite being in a more multimodal one (Kroesen, 2014).

Concerning the ownership or the number of cars, having more cars is associated with lower multimodality measurements and/or being in more monomodal (car-based) user groups (e.g., De Haas et al., 2018; Heinen & Mattioli, 2019a; Susilo & Axhausen, 2014). Likewise, the opposite also holds, that individuals are more likely in more multimodal classes or classes where the car is complemented with other modes when having fewer cars in general (e.g., De Haas et al., 2018; Diana & Mokhtarian, 2009;

Kroesen, 2014). The preceding aligns mostly with findings on whether to own a car (e.g., An et al., 2022; Ton et al., 2020).

Generally speaking, car(licensure)-dependency is visible in earlier studies, as individuals with car (licensure) ownership mainly show non-multimodal behaviour. However, despite the known car dependency, the effect of company cars, besides the other mentioned resources, is not yet included in the mentioned research, which was also an earlier identified knowledge gap (section 1.2).

Bicycle ownership

Some studies included bicycle ownership when analysing the relation with multimodality. This is now explained and discussed at the end.

Two studies show that higher bicycle ownership of individuals means higher multimodality measurements of their travel patterns (Heinen & Chatterjee, 2015; Heinen & Mattioli, 2019a). Also, when multimodality is measured separately for work, maintenance and leisure trips, a higher level is seen for those owning a bicycle (An et al., 2023). Furthermore, when owning a bicycle, individuals are also more likely not to be car owners and to be in not-so-multimodal bicycle-exclusive classes. However, they are also more likely to be in more multimodal public transport-oriented classes (Ton et al., 2020), with a higher multimodality measurement in the study of An et al. (2022).

The previous showed that a limited amount of studies included bicycle ownership, which researchers regard as important for defining multimodality (section 1.2). Moreover, only a few studies with the Dutch scope involved this variable, while The Netherlands (our scope) is regarded as an active travel-oriented country (section 1.7.2), which makes it interesting to analyse bicycle ownership further.

3.3.3 Built-environments variables

Characteristics of the built environment are researched in several studies concerning multimodal behaviour, generally for lower and higher densities. After this comparison, the results of this variable are concluded.

For example, regarding urban (population) density, it is found that more rural settlements compared to inner-city living negatively affects the multimodality measurement of travellers (Heinen & Chatterjee, 2015; Heinen & Mattioli, 2019a; Susilo & Axhausen, 2014). Furthermore, many individuals living in rural areas are mainly in unimodal strict car user groups (An et al., 2022; De Haas et al., 2018; Schneider et al., 2021; Ton et al., 2020; Olafsson et al., 2016).

On the other hand, a greater multimodality measurement is found for individuals with a higher residential land-use mix (i.e. degree of balanced land-use purposes) for several trip purposes such as work and maintenance (An et al., 2023). Also, when living in more urban or populated (dense) areas, a higher likelihood exists to be in more multimodal travel classes (De Haas et al., 2018; Kroesen, 2014; Olafsson et al., 2016). Furthermore, more multimodality is reflected in some classes in which public transport use plays a major role, where those classes are also characterised by individuals living in highly urban areas (An et al., 2022; Schneider et al., 2021; Ton et al., 2020). However, more unimodal bicycle-oriented groups also have a higher share of individuals living in urban areas (Ton et al., 2020; Olafsson et al., 2016).

The previous shows that the urban density is operationalised very differently among research, while all show overall an effect of more multimodality in more urban areas, which makes it essential to include this variable as part of the framework for analysing multimodal travel behaviour. This framework, with chosen variables and operationalisation (based on all the previous), is explained in Chapter 4.

4. Methodology

This chapter elaborates on the methodology for executing the main research. First, the method of LCA is explained (section 4.1). Second, based on the LCA explanation and the previously identified theoretical background (Chapters 2 and 3), the model is conceptualised (section 4.2). Third, the used data is explained (section 4.3), and fourth, the travel behaviour is defined (section 4.4) in order to operationalise the data as a fifth step (section 4.5). Sixth, the dataset is further processed (section 4.6), and seventh, the final measurements are created (section 4.7). The descriptive statistics of the final sample are shown (section 4.8). Last, the modelling strategy to decide upon the number of classes for the latent class model is explained (section 4.9).

4.1 The LCA method

This section dives deeper into the research method chosen (section 1.7.2) with the characteristics (latent class analysis and latent class analysis vs cluster analysis), the suitability and the limitations of the chosen method Latent Class Analysis (LCA). The method is applied to estimate and determine the number of classes, as shown in section 4.9.

Latent Class Analysis

As explained by Weller et al. (2020), LCA is a statistical procedure, also known as Latent Class Cluster Analysis (LCCA). This procedure can identify different subgroups, called classes or clusters, within a population which are qualitatively different. For instance, a named bicycle class or a public transport class. These subpopulations are not observed or organised groups in real life, meaning they are called latent classes of cases. So, the characteristic cannot be observed directly (Vermunt, 2004), i.e. in our case, being multimodal or not is not directly observable when you 'ask' someone. LCA is thus also known as a mixture model, where hidden groups can be identified (Oberski, 2016). According to Weller et al. (2020), cases can be individuals who are grouped based on comparable patterns.

The assumption of LCA is that several latent classes exist in society (Lee et al., 2020). A latent class can also be seen as a modality style, where the behavioural predisposition of a person (reflected by a particular style) determines the use of several means (Vij et al., 2013). In LCA, patterns of observable survey questions or scores, for instance, are assumed to explain class membership. More specifically, in modal travel behaviour research, classes are characterised by relative homogenous mode use patterns while maximising heterogeneity of mode use patterns between classes (Lee et al., 2020; Lezhnina & Kismihók, 2022). These observable variables are called indicators (Vermunt & Magidson, 2002). Moreover, Weller et al. (2020) explain that outward characteristics could be shared among people as class members. Outward characteristics are called covariates and are not 'part' of the model. However, these shared characteristics describe the people in the obtained groups based on the indicators. Because of the inclusion of covariates, predictions can also be made for individuals not part of the model.

LCA versus cluster analysis

After LCA is briefly explained, it is useful to explain both the general and more specific clustering methods and the differences, where the advantages of LCA (advanced clustering method) become clear over traditional clustering. Clustering is an unsupervised learning method, meaning that cluster classifications are driven by the data used (Sinha et al., 2021). Both LCA and traditional cluster analysis are person-oriented analyses (Collins & Lanza, 2010, in Weller et al., 2020), where patterns across cases are used to group individuals in classes, as opposed to the variable-centered approaches, where the relationships between variables are examined. However, the latter is also possible in cluster analyses. Statistical and theoretical criteria are used to determine which solution is best for the series of solutions in LCA and clustering (Weller et al., 2020). For LCA, the series consists of solutions where every solution has one class more than the previous solution, whereas traditional clustering usually has pre-determined number of classes. For both methods, multiple variables can be used to cluster cases and identify patterns, which is useful for our research aim of using several variables measuring the modal usages. Moreover, the extent of multimodality can be calculated for the identified clusters in both methods. However, the advantage of LCA is that it is possible to predict class membership based on the mentioned covariates and indicator values.

Several more distinctions can be made to understand how LCA, one of the clustering techniques, works and differs from the more traditional clustering techniques, generally referred to as k-means clustering.

First, the assumption for clustering (read: k-means clustering) is that the most comparable scores across variables of interest belong to the same cluster. In contrast, LCA posits that society already has existing latent classes, which explicates patterns composed of comparable observed scores across cases, and between classes, the patterns are heterogenous (Weller et al., 2020). However, individuals are not 'assigned' to classes in LCA, as the model generates probabilities for class membership for all classes (Sinha et al., 2021).

Second, clustering is a distance-based method which separates observations in a specific cluster based on a dissimilarity criterion to identify homogeneous clusters. On the contrary, LCA is an advanced probabilistic model-based clustering method where probabilistic models are fitted to the data, which reduces misclassification biases (Araghi et al., 2017; Lezhnina & Kismihók, 2022). Each case can be assumed to belong to one class, but the uncertainty about class membership is considered by having different probabilities of belonging to several classes (Vermunt & Magidson, 2002). This allows cases not included in the model to calculate the posterior class-membership probabilities based on the estimated model parameters and the observed scores of cases. However, both procedures can generate categorical classification variables to be used in other analyses (Weller et al., 2020). For both methods, the criteria to choose the right amount of classes differ because of being probabilistic in nature or not. To specify the working further, in short, both algorithms of the methods are explained in two points.

Clustering is based on a pre-determined amount of clusters. For instance, two random cluster centres are appointed when clustering cases with data on two variables in two clusters. Based on the nearness of those case values (in the x-y space) to those centres, the cases are assigned to the nearest cluster. Then, the cluster centre is recalculated based on the variable means, and cases are reappointed to one of the two clusters with the nearest cluster centres. This continuously iterates until the cluster centres remain relatively stable (Shukla & Naganna, 2014) and everyone is appointed to one specific cluster. This process results in different solutions when using the algorithm because of different start values. Only a few and less strict (statistical) tests can guide in choosing the right amount of clusters (Magidson & Vermunt, 2002).

Maximum Likelihood algorithms obtain the latent class models, and the effect of covariates on class membership is obtained using multinomial logit models (Vermunt & Magidson, 2016). In LCA, the use of different start seeds is automated in the software package, Latent GOLD, to prevent acquiring a local (and very different) model solution (Vermunt & Magidson, 2016). The algorithm is extremely stable, but slight parameter differences can occur between several runs. As several solutions can be acquired by using several classes, it should be decided upon the number of classes. Various rigorous statistical tests guide choosing an appropriate number (Magidson & Vermunt, 2002).

Third, the data scale of variables is continuous for clustering analysis because the variable's means are used to define the most comparable scores (Weller et al., 2020). Compared to clustering, LCA uses mainly categorical measured indicators, which is a traditional LC model (Lezhnina & Kismihók, 2022; Magidson & Vermunt, 2002). However, a mixture of data scales is possible: continuous, categorical (nominal or ordinal), or counts scales, and combinations of them (Magidson & Vermunt, 2002).

LCA suitability

As identified in section 1.5, LCA is suitable for analysing multimodal travel patterns. The method is more in-depth explained, and LCA is deemed more appropriate than the general clustering technique based on the previously mentioned aspects. However, it should be acknowledged that multiple methods, besides clustering (quantitative), could be suitable. Although qualitative research is used, these analyses are regarded as subordinate to the main quantitative method, so only LCA is compared.

The advantage of LCA is that a contextualised understanding of travel behaviour can be acquired, in which subgroups of individuals can be identified (Weller et al., 2020). So, this quantitative method's holistic view of travel behaviour is also advantageous compared to other quantitative methods like regression (see below), where the average traveller is captured instead of LCA with more specific travel user groups. This limits the general downside of quantitative research compared to qualitative research, which generally has a better view of the context and setting of people (Creswell, 2009, p. 16, 17).

LCA is more suitable when comparing LCA with more specific other (basic) quantitative or qualitative methods, for instance, when identifying the potential combinations of travel modes which (scarcely) exist by using interviews, literature studies or descriptive statistics. Because of the quantitative characteristics of our method, a parsimonious model can be found with the latent class clustering algorithm. In other words, a model with the smallest amount of latent classes (travel user groups), which still describes the associations between indicators well (Kroesen, 2019), can be found by identifying only the combinations of travel modes which exist in reality. Furthermore, the intrapersonal variability (section 1.2) used in our research is suited to be analysed with LCA (a segmentation technique). As explained, various patterns are reduced to smaller clusters, including individual patterns (Ton et al., 2019).

Moreover, LCA has the advantage of effectively including combinations of mode usage, as several variables can be used to determine the classes. In contrast, other methods (like regression) are more suitable for including a dependent variable which captures one travel mode or dependent variable. Regression analysis (quantitative) is useful for identifying the effect of independent variables on the dependent variable by mutually checking the independent variables for overlap in effect, as explained by Chatterjee and Hadi (1991, p. 1, 58). LCA also allows for including individual effects of variables on latent class membership, as explained.

LCA limitations

Nevertheless, the advantages and suitability aspects of LCA, every method or viewpoint towards travel behaviour has inherent limitations. These are in this paragraph addressed. Firstly, because cases are assigned to classes based on their probability, the proper class assignment is not assured, and the exact number of sample members cannot be identified (Weller et al., 2020). This can be overcome by using a large dataset, limiting the impact of misclassified individuals. Next, qualitative names are mostly assigned to classes because of the complexity of model parameters, which may lead to a 'naming fallacy' if the names are not representative enough (Weller et al., 2020). The naming fallacy could be reduced by emphasising interpretability and communicating the characteristics of the classes. Last, another downside is that there is no optimal (statistical) criteria set for comparing solutions with different amounts of classes, making the approach (statistical) exploratory (Weller et al., 2020). The latter downside can be assessed by explaining which criteria (including statistical and interpretability) are used for comparing the latent class solutions and which considerations are made.

4.2 The model conceptualisation

Based on LCA (section 4.1), the model is set up (Figure 7). First, the conceptual working is explained.

An LCA model consists of a measurement model with indicators and a structural model with covariates. The latent classes to be determined are nominal, and the indicators are categorical (nominal or ordinal), continuous or count data scales. The covariates are mainly categorical (nominal) or continuous (Vermunt & Magidson, 2016). The measurement model describes how the latent variables are measured and explain all the indicators' variations. The structural model describes the relations between the covariates and the latent variables. Inactive covariates can describe the different latent classes and do not affect the model, as De Haas et al. (2022) explained. The individuals are assigned to a latent class with the highest latent class membership probability (Vermunt & Magidson, 2016). Nevertheless, all the outcomes are still based on all the individual probability values.

The causal relations are as follows. The assumption is that covariates causally precede the latent variable (X), whereas the latent class variable causally precedes the indicators. So, the latent classes affect the indicator values, where the covariates are assumed to have a causal relation with the latent classes. The discrete latent variable (X) (with classes) can account for the observed associations (i.e. correlations) between the indicators (Magidson & Vermunt, 2004). It implies that LCA assumes that local independence exists, i.e., insignificant associations between the indicator variables, conditional on the latent variable. This tends to result in classes where variables are unrelated to each other within each class so that observations are similar to each other but different from those in other classes (Oberski, 2016). For the covariates, it is possible to see how these are related and affect class membership. The assumption is that covariates are not influencing each other (powerfully). Moreover, the active covariates are controlling for each other, meaning that individual effects of individual covariates can be assessed.

The variable (groups) from the conceptual framework in Figure 7 is explained. Based on the previous literature reviews (Chapters 1, 2 and 3), the active covariates are socio-demographic variables, mobility resource variables, built-environment variables and year. The multimodality measurements and residential municipality are added as inactive covariates to describe the emerged classes. As explained, the multimodality measurements added to the latent classes can describe the extent of (quantitative) multimodality for several multimodal and modality styles (Diana & Pirra, 2016). The multimodality measurements are thus not affecting the model, which is preferred, as they measure (partly) the same as the indicators. The residential municipality is only used for showing the distribution and not for using it as a determinant, which would also result in too many parameters to be estimated. Moreover, indicators about mode usage are included in the model (measuring travel behaviour).

The specifically used variables based on this conceptual framework for the covariates are explained from the data operationalisation onwards (section 4.5), and the correlations of covariates are checked for in section 4.8 (descriptive statistics). The indicators consist of mode usage (travel behaviour). Their specific intensity measuring and operationalisation/definitions are discussed in sections 4.5/4.7. We use four modes: walk, bicycle, public transport (bus, tram, metro and train) and car (passenger and driver) for the measurement model. Although public transport both involves interand intra-city travel and could be split into bus/tram/metro and train to capture their different intensity of use better (e.g., Molin et al., 2016), this operationalisation is used more often in other studies (e.g., De Haas et al., 2018; Schneider et al., 2021; Ton et al., 2020). Besides, using these four modes (categories), it is assumed that all modes are represented to some extent, and it does include active travel modes walk and bicycle (identified as important).

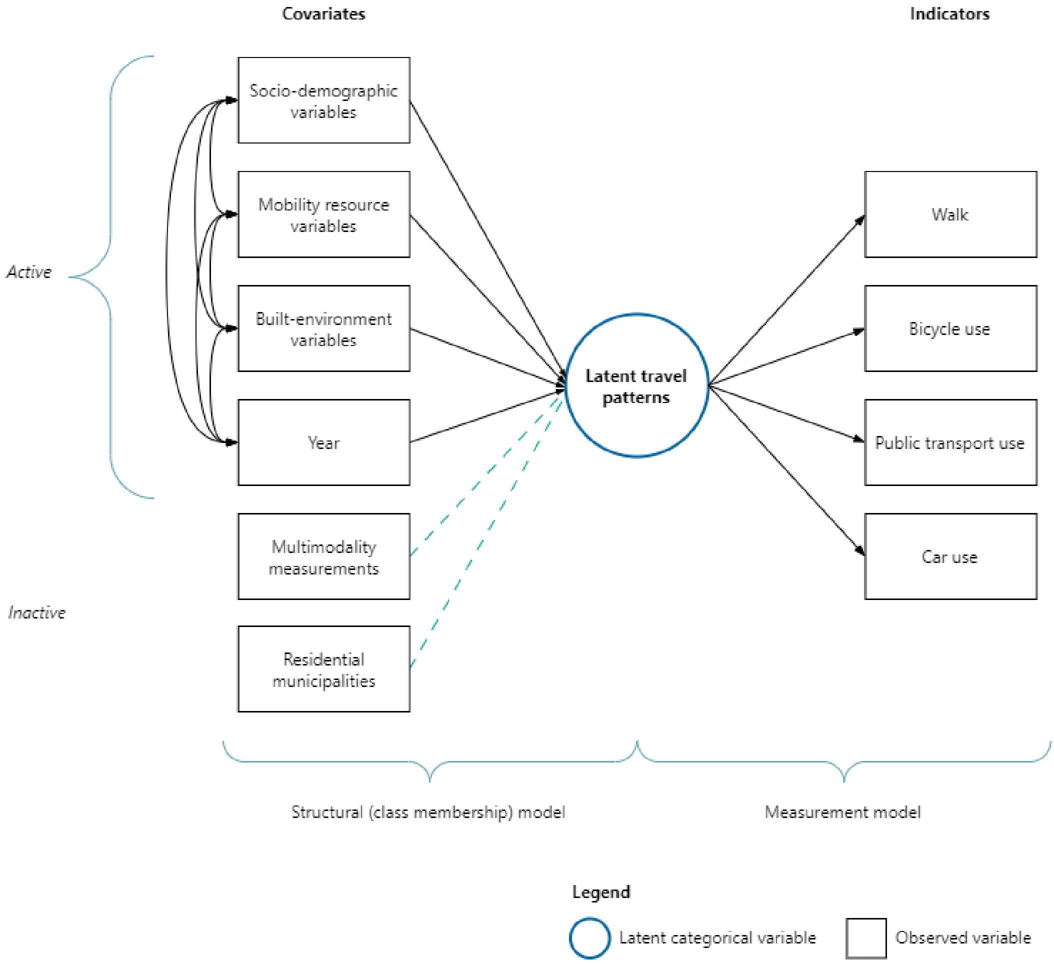


Figure 7. Conceptual framework studying multimodal travel patterns (layout adapted from Molin et al. (2016)).

Some notions about how the variables are included, given before, are explained.

The mobility resources are likely endogenous to travel patterns. This implies that, due to self-selection, someone that likes to have a specific mobility style is likely to acquire the several modes to be used (the other way around than assumed and specified in the model). So, owning several mobility options cannot be regarded entirely as exogenous in determining the latent class with respecting modal patterns. However, these mobility resources are still included as active covariates (instead of using them only to describe the obtained classes) to acquire some information about the strength of those effects, which should be interpreted with care.

The year variable could have been used as an inactive variable, as it can be used to describe the classes without being part of and affecting the model. However, it is incorporated as an active covariate, entirely part of the model, and assumed to be exogenous on the latent classes. This is especially less problematic when the correlations between the years and the others are low (section 4.8). The classes were once identified in LCA based on the pooled data for several years. By including the year variable, the sizes of all the identified classes per year can be identified.

4.3 Data and sample

This section explains the used data for the conceptual model (section 4.2) and their implications and the defined sample population.

Data

In order to create the latent class model (section 4.2), this study uses Dutch data (see scope in section 1.7.2). Data is used from a National Travel Survey (NTS), called the Dutch National Travel Survey (OViN, in Dutch: Onderzoek Verplaatsingen in Nederland); see Statistics Netherlands & Department of Waterways and Public Works (2010-2017). The OViN is executed nationally to examine the daily travel patterns of individual Dutch residents. The OViN is administered by Statistics Netherlands (in Dutch: Centraal Bureau voor de Statistiek (CBS)) for the Department of Waterways and Public Works (in Dutch: Rijkswaterstaat (RWS)), which is the executive agency of the Dutch Ministry of Infrastructure and Water Management. The data is accessible on request on the DANS Data Station Social Sciences and Humanities. Other used data of Statistics Netherlands is accessible on CBS StatLine.

Since 1978, CBS started and is administered for researching mobility cross-sectionally continuously, as described in the research setup (Statistics Netherlands, 2018). Since then, the research setup and responsible administration have changed significantly once in a while. The survey used for our study, the OViN, was executed by CBS from 2010 to 2017 with a comparable research setup. From 2018 onwards, the research setup has changed and continued under another name, ODiN (Onderweg in Nederland). The preceding means that the mobility patterns of OViN cannot be compared with earlier or later investigations. It is noteworthy that OViN is less recent than ODiN but has more completed years to show trends.

OViN uses many representative individual respondents. The number of respondents is about 40,000 individuals per year, about 0,2% of the target population. Every year, the random sampling determines which individual Dutch residents will receive an invitation to complete the survey for one day. Travel information is collected for every day in a year from January to December by assuring that every day has an appropriate response amount. It is also accounted for to have an optimal national spread of respondents compared to regional population sizes. The distributed surveys are filled in online and, in case of no response, potentially by telephone or face-to-face. For one day, people are asked to record and fill in a travel diary about the trips made, including, among other things, the origin and destination, trip purpose, mode(s) of transport, start and end times and travel distances. In addition, personal and household characteristics, such as licensure and ownership of modes, are also filled in the survey.

The OViN has advantageous characteristics, also compared to comparable surveys of other countries or other surveys in The Netherlands. OViN is the only yearly repeated high-quality survey in The Netherlands with a representative large set. It includes various modes, making it suitable for showing multimodal trends of a population. The cross-sectional nature means that a representative (different) population set can describe changes across the population every year. This way, the trend can be viewed comprehensively by employing a national travel survey (An et al., 2021). The OViN is now compared with other surveys.

When comparing the OViN with the 7-day diary of the NTS in England, for instance, where day-to-day variations or a weekly pattern can be captured, diary fatigue is less of an issue. Diary fatigue, resulting in fewer accurate reported trips and lower quality, is less of an issue in shorter surveys (An et al., 2022), like the one-day travel diary of OViN.

The OViN can also be compared with another travel survey in The Netherlands with a longer travel diary, The Netherlands Mobility Panel (MPN), which commenced in 2013 by longitudinally tracking mobility behaviour. Although a longitudinal survey can capture changes within individuals, predictors of changes, or (potential) causal relationships (Heinen & Mattioli, 2019a), longitudinal surveys often use fewer respondents. Buehler & Hamre (2015) also identified that longer travel diaries mostly have relatively small sample sizes due to collection efforts. In the case of MPN, a longitudinal household panel, it uses a 3-day travel diary but consists of only 2000 households, and it has limited information about active travel modes compared to NTS (De Haas et al., 2018; Ton et al., 2020). Furthermore, in the cross-sectional research of OViN, compared to MPN, no issues arise from the cohort effect when respondents get older or have life events which change their behaviour (De Haas et al., 2022).

In order to visualise the outcomes on the municipal level (inactive covariate in the conceptual model in section 4.2), digital geometry data of municipality boundaries is used. Data from 2017 from the boundaries from The Netherlands' Cadastre, Land Registry and Mapping Agency (in Dutch: Kadaster). This data is published by CBS (The Netherlands' Cadastre, Land Registry and Mapping Agency & Statistics Netherlands, 2019) on the CBS file geographical data. The columns 'municipality name' (in Dutch: GM_NAAM) and geometry (Geometry) are used. The data for merging municipality names from 2010-2017 to the 2017 municipality organisations to visualise them for the most recent year is based on the published data of CBS and a created list of previous Dutch municipalities (Wikipedia, 2023).

Using the 2017 municipal division results in the municipality names being changed for the years before the reorganisation if the residential municipality (of an individual) was part of a reorganisation at some point in 2010-2017. This implies that the residential municipality map does not precisely correspond with earlier years, which is yet challenging to see by eye. In order to keep the correct urban density of individuals in previous years before the reorganisation, the urban density has not changed to the urban density of the municipality after the reorganisation. In this way, the original urban density of an individual in that year is still accurately reflected. Still, it could be that another individual, after the reorganisation in a later year, lives in the same city but has a different urban density because the municipality (and urban density) has become more or less dense. The own city itself could have remained at the same urban density, but now it is part of a municipality involving more cities or suburbs. Besides the reorganisations, municipalities could have grown when new addresses were built in the meantime, or they could have even shrunk when areas were demolished. Thus, irrespective of using the same municipal division for all the years, the case of changing urban densities would have always been the case. However, it has limited impact and creates more recent and accurate results about the municipalities.

Sample

The used OViN data has information from the population from age 0 onwards. Previous studies included people in their sample from being an adolescent (about 10-20 years old) onwards (e.g., Scheiner et al., 2016; Ton et al., 2020). Others include people in their sample who were adults or older (e.g., An et al., 2022). Our research defines the sample as people being 18 or older (adults). This is because children typically have different travel patterns in several parts of their lives (different education types) and when they become an adult. Also, as fresh adults, they can potentially drive themselves, for instance. Moreover, most children likely use mainly the mode bicycle, either or are being brought (by car), and are less likely to travel daily by public transportation. At the same time, mode diversity is important for multimodal travel behaviour. So, in this research, the modality styles are only explored for individual adults in The Netherlands to be more potentially generalisable for a broader and comparable population (adults only).

4.4 Definition of travel behaviour

At the data-operationalising, processing and measuring stage (sections 4.5, 4.6 and 4.7), it is important to know how to measure travel behaviour. The operationalisation is influenced by the data used (section 4.3). Multiple considerations are taken into account to define travel behaviour: general aspects, level of analysis, the measure of intensity, the mode set and the time period.

Travel behaviour: general aspects

The first general aspect of travel behaviour (or the scope) is location. This research only analyses travel behaviour for travel in The Netherlands (section 4.3). The results represent the mobility patterns on Dutch ground (without foreign travel). The second general aspect is immobility. People who do not travel are not taken into account. By excluding the immobile people on the day of the diary, no one can be misclassified in a potential light traveller group. The last general aspect is the type of travel. Only day-to-day or daily travel behaviour is considered. One can think about commuter traffic or recreational purposes. These travel purposes reflect the 'normal' behaviour best compared to, for example, holiday travel or travelling for foreign work purposes. Then the question arises of how to deal with people going on holiday (non-daily travel) after doing the groceries (daily travel), for example. Excluding the part of people's captured patterns that are not marked as daily travel (e.g., holiday or foreign travel) would result in incomplete mobility patterns, thus biasing the amount of travel intensity and their modal patterns. Therefore, this research only includes individuals if their full travel pattern comprises daily travel.

Travel behaviour: level of analysis

Travel behaviour can be captured at the trip or stage level. A trip comprises stages (part of a trip). In the research of OViN, both levels are captured. An imaginary example of a travel pattern in a certain timespan to explain the trip (in Dutch: verplaatsing) or stage (in Dutch: rit) level is shown in Figure 8. This travel pattern example comprises two trips. Trip one has three stages involving all other modes, with the origin at location A and the destination at location B. The bicycle is the access mode for the mode train, and the egress mode is walking to arrive at the destination. For instance, an activity (not further specified) can be performed at that location. At a later time, at origin B, this individual made the second trip consisting of one stage by car to the end-point C. This example assumed that the car was available at point B, irrespective of not being the start location. For this trip, the purpose was, for example, going on a visit. In the OViN data, the destination purpose of a trip can be, for instance, going on a visit/stay, going to work, or going home. The (higher level) motive purpose of a trip can correspondingly be going on a visit/stay, from and to work, and going home. The latter can be the case after trips with all potential destination purposes. When the purpose is going for a drive/walk, it is accounted as one trip with one stage in OViN (as it has the same origin/destination). It should be noted that the short walking trip between leaving the bicycle to the train station/platform is not taken into account, and the short walking trip after parking the car to the destination is also not separately tracked in OViN (these are transfer stages).

Travel behaviour can thus be measured on the trip or stage level. As Chapter 2 shows, multimodality measures can also be measured on the trip or stage level. This study proposes to use the stage level, in line with studies involving multimodality measures, which mostly (only) used indicators on the stage level (e.g., An et al., 2022, 2023; Heinen & Chatterjee, 2015; Heinen & Mattioli, 2019a). Stages as the unit of analysis allow for taking the secondary (or tertiary, et cetera) modes of a trip into account (Heinen & Chatterjee, 2015) instead of looking at the primary mode of a trip or the combination of modes per trip. As explained by Heinen (2018), the trip level accounts for variation in the (combinational) mode used between trips, whereas the stage level accounts for variation in all possible modes used, irrespective of the mode combinations used and the length of each stage, where often the lengthiest stage is marked as the main mode of a trip. The stage level accounts thus for variation among the entire travel pattern. In the example of Figure 8, the main mode of trip one would be the train, whereas including all the modes reflects the multimodality within and between trips.

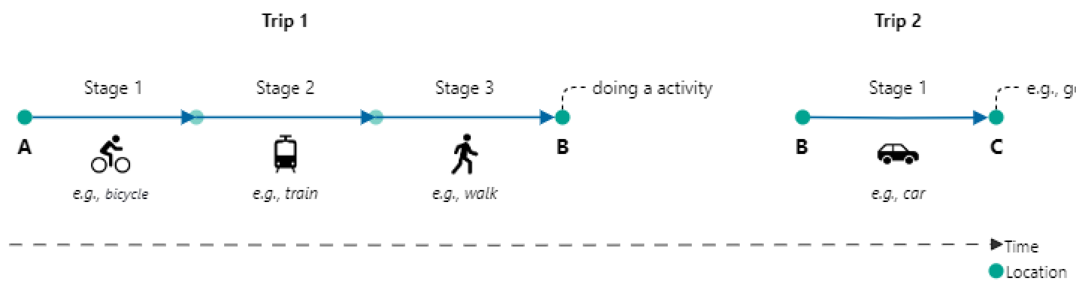


Figure 8. Imaginary example showing the trip-stage levels in travel behaviour.

Travel behaviour: the measure of intensity

The travel behaviour (on the trip and stage level) can have different measures of travel intensity per travel mode. As said earlier, measures of travel intensity are mainly the number of trips/stages, travel times or travel distances (e.g., Diana & Pirra, 2016). Due to the self-reported travel distances and times in the OViN study, the number of trips (in our case stages) is regarded as the most reliable (De Haas et al., 2018). Travel distances and times are likely to have a skewed distribution, which can be overcome by creating ordinal groups. However, as De Haas et al. (2018) explained, trips are count variables, which can be approximated with a Poisson distribution in a latent class model.

The advantage of using the stage (or trip) rates per mode becomes evident after further explaining the disadvantages of using the travel distances and times. Ton et al. (2020) explain that the use of distance leads to an overrepresentation of the car and public transportation intensity as these are more attractive modes for longer distances. Moreover, travel distances are also more suited to analyse the relationship between travel behaviour or multimodality and the environmental impact (Diana & Pirra, 2016). Concerning travel times, it shows a wide range of intensity for public transportation use due to inter- and intra-city travel (Ton et al., 2020). Moreover, travel means have different speeds. The different means are thus used for trips with different frequencies, like systematic or recreational trips (Diana & Pirra, 2016), making travel times less representative of the intensity of the use of specific modes.

Travel behaviour: mode set

Moreover, the number of modes included in the mode choice set must be considered, as explained in section 4.2. It is explained by Heinen and Mattioli (2019a) that a person using three modes in a balanced way can be measured as very multimodal when using that three modes in defining mobility patterns, compared to using eight modes, for instance. However, in studies using 3 and 8 modes with multimodality measures, the overall trends or results have some different values but have generally similar results (e.g., Heinen & Chatterjee, 2015; Heinen & Mattioli, 2019). We use four modes: walk, bicycle, public transport (bus, tram, metro and train) and car (passenger and driver) (section 4.2). Again as for the daily travel, if someone used another mode and only part of the travel pattern is deleted, someone might look less multimodal, as fewer modes and the travel intensities do not reflect the whole travel pattern well.

Travel behaviour: time period

Last, the timespan of the travel patterns can be reflected upon. As mentioned, the OViN is a one-day diary, relatively short compared to weeklong or multi-week cross-sectional studies (Buehler & Hamre, 2015). It is argued among reviewed research that surveys of one week capture the typical variability in everyday habitual travel best, compared to multi-week studies, which capture additionally occasional travel behaviour. Moreover, the longer the time period of a survey, the more likely an individual uses more modes of transport (Nobis, 2007). Although the OViN has advantages compared to other surveys (section 4.3), the implication of using a one-day diary is the potential misclassification of individuals as only a part of their weekly patterns is captured, and only a part of their potential multimodal behaviour is captured. This limitation is partly overcome by having a large sample set, resulting in a higher likelihood that many individuals were having a representative day on the day of the travel diary task.

After all, the operationalisation choices of the sample and travel behaviour can be summarised in the following:

Travel behaviour operationalisation (summary)

Travel behaviour is measured for adults doing 'daily travel' in The Netherlands on the stage level (stages are part of trips), measured in the number of stages as a measure of intensity for the modes walk, bicycle, public transport (bus, tram, metro, train), and car (driver and passenger) for the period of one day.

4.5 Data operationalisation

This section describes the operationalisation of variables (in values or categories) based on the OViN data (section 4.3). First, Table 8 shows the overview of used variables per variable group. Variables are included based on the conceptual model, defined travel intensities and mode set (indicators), literature review on multimodality measurements (inactive covariates number of modes and OM_PI), literature review on determinants (active covariates), the knowledge gaps (year, residential municipality), and definition of multimodality, explained in earlier sections or chapters. In the following sections, some original OViN variable names are mentioned. It should be noted that for some variables used in this section, the name was slightly different in the dataset in some years, mainly because sometimes a capital letter (or not) was used in the name. It is helpful to know when using syntax to work with the data.

Table 8. The conceptual model specified in variables.

Overarching variable group	Variable groups	Variable
Indicators		
Travel behaviour	Travel behaviour (over one day)	# Walk stages
		# Bicycle stages
		# PT stages
		# Car stages
Active covariates		
Socio-demographic variables	Personal characteristics	Gender
		Age
		Ethnicity
		Education level
	Household status	Household composition

	Employment status	Occupation
	Economic status	Household income
Mobility resource variables	Mobility resources	Licensure
		Ownership household car
		Ownership company car
		Ownership household bicycle
Built-environment variables	Built-environment characteristics	Urban density
Time	Time	Year
Inactive covariates		
Travel behaviour	Multimodality measurements (over one day)	Number of modes
		OM_PI*
Built-environment variables	Built-environment characteristics	Residential municipality

*Objective Mobility Personal Index. A Higher value indicates a higher degree of multimodality.

The following sections describe the operationalisation to be used for the indicators and the operationalisation for the active covariates and the inactive covariate (residential municipality). The measurements of travel behaviour (indicators) and inactive covariates (multimodality measurements) are described in section 4.7, after the description of the data filtering, aggregation and created final sample.

Indicators

As the indicators are about the number of stages per mode (Table 8), the data needs to be operationalised shown in Table 9, to calculate this measure (section 4.7). The original OViN variable name is given (in Dutch), and the description of the original variable. The created categories and the original OViN categories are shown with the original coding. A variable measuring the stage mode (OViN variable, in Dutch: Rvm), which records the used mode on the stage level, is transformed into the four specified mode categories walk, bicycle, PT (bus, tram, metro, train) and car (driver and passenger). These categories on this 'Rvm' variable can be used later for the measures of travel behaviour.

Table 9. Operationalised variables (to be used for indicators) on the stage level.

OViN variable (in Dutch)	OViN variable description	Created categories	OViN category (code)
Indicators (used for)			
Rvm	**Stage mode**: captures per stage which mode has been used.	**Walk**	On foot (22)
			Pram (23)*a
		Bicycle	Bicycle (electric and/or non-electric (15)
			Bicycle as passenger (16)*b
		PT	Bus (only PT) (5)
			Tram (4)
			Metro (3)
			Train (1)
		Car	Car driver (6)
			Car passenger (10)

*Prevalence in final sample: < 0.00% of total stages.
a: For adults, most likely measuring active travel (pushing a pram on foot).
b: Not measuring active travel.

Next to the explained modes per category, some not mentioned categories are also used in our four mode categories, as seen in Table 9. The created walking category consists of the OViN modes on foot and pram. As pram use is marked as on foot in the main trip mode variable (OViN variable, in Dutch: Hvm), this is also included in the created category walk. Most likely, it is used as an adult to push it, as otherwise, it could have been marked as disabled or in the other mode category. However, the prevalence is remarkably low. An implication for the created category bicycle and the used OViN category is that the electric bicycle is also recorded in the bicycle mode category (from 2013 onwards, this information can be split up, which is not used in our research). An implication is that this analysis presents someone using an electric bicycle as a somewhat more active traveller. Moreover, the bicycle as a passenger mode category (not measuring active travel) also belongs to the main bicycle category in OViN and our research. However, the prevalence is also really low, thus having a low impact on the results.

Active covariates

According to Table 8, the chosen variables for the active covariates are also operationalised based on the original OViN categories in Table 10. This reduces the number of categories and better interprets the outcomes. The original OViN variable name (which might be an abbreviation) is given (in Dutch), and where necessary, the description of the original variable. The created variables can be described by their name (in bold) and the transformed categories used. The original OViN categories, with the original coding, are given to show how these are transformed into newly created variables. When the categories remain the same as in OViN, only the code number is given, or for continuous variables, nothing is given (age and year). The categories involving the unknown or other category are mostly transformed into a missing value. For the mobility resource ownership variables, the unknown category is mostly transformed into the no category because of those variables' yes/no structure, and the prevalence of the unknown category is relatively low.

The categories which remained the same can be seen in Table 10. These are gender, age, education level, household income, (car) licensure, and year. The other categories are explained and grouped per variable group (see Table 6).

People who are Western and non-Western immigrants - (non-)Western is based on their birth countries - have been merged into the immigrant category. The native category remained the same. The household composition remained the same for the categories one-person and couple households. The merged other category can also consist of two persons, but this is not identifiable due to the original category setup. Moreover, the other category mainly consists of three or more persons when looking at the original categories. The original other composition category (8) can, for instance, consist of only 'other' people, like brothers/sisters, friends or students (who are not living in a one- or couple household). For occupational status, people working 12-30 hours a week and people working 30+ hours a week are used for the employed category. Pupils/students are used for the category student, and the other category consists of people who are working in the household, unemployed, unable to do work or retired.

The ownership of a household car (yes/no) is based on a variable measuring the number of cars in the household. The household car, instead of the individual ownership, is chosen, as the household car likely better represents the ability to use that resource. The ownership of a company car (yes/no) is based on the variable measuring the ascription/registration of the car of the main car user, in which it is reported if the ascription is on a company. The ownership of a household bicycle is based on the variable measuring the number of non-e-bicycles in the household. Based on a textual/format change in the question about bicycle ownership and added questions about e-bicycle ownership in

Table 10. Operationalised variables (active covariates) on the individual level.

OViN variable (in Dutch)	OViN variable description	Created variable Transformed categories	OViN category (code)
Active covariates			
Geslacht	-	**Gender**	
		Male	(1)
		Female	(2)
Leeftijd	-	**Age (18+)**	
Herkomst	Native: both parents are born in the NL. Immigrant: else.	**Ethnicity**	
		Native	Native (1)
		Immigrant	Western immigrant (2)
			Non-Western immigrant (3)
		Missing	Unknown (4)
Opleiding	Highest level of completed education.	**Education level**	
		No education	(0)
		Primary education	(1)
		Pre-vocational education	(2)
		Vocational & higher secondary education	(3)
		Higher vocational & university education	(4)
		Missing	Other education (5)
			Unknown (6)
HHSam	-	**Household composition**	
		One-person household	(1)
		Couple household	(2)
		Other household	Couple + child(ren) (3)
			Couple + child(ren) + other(s) (4)
			Couple + other(s) (5)
			Single parent family + child(ren) (6)
			Single parent family + child(ren) + other(s) (7)
			Other composition (8)
MaatsPart	Social participation.	**Occupation**	
		Employed	Employed 12-30 h/week (1)
			Employed ≥ 30 h/week (2)
		Student	Pupil/student (4)
		Other occupation	Own household (3)
			Unemployed (5)
			Incapacitated (6)
			Retired (7)
		Missing	Other (8)
			Unknown (9)

Table 10. (continued). Operationalised variables (active covariates) on the individual level.

OViN variable (in Dutch)	OViN variable description	Created variable Transformed categories	OViN category (code/value)
HHGestInk	Standardised disposable yearly household income (corrected for household size and composition).	**Household income (standardised)**	
		< €10,000	(1)
		€10,000 - €20,000	(2)
		€20,000 - €30,000	(3)
		€30,000 - €40,000	(4)
		€40,000 - €50,000	(5)
		≥ €50,000	(6)
		Missing	Unknown (7)
Rijbewijs	Car licensure.	**Licensure**	
		No	(0)
		Yes	(1)
		Missing	Unknown (2)
HHAuto	Number of cars in the household.	**Ownership household car** (i.e. having one available in the household)	
		No	0
			Unknown (10)
		Yes	1..8
			9 cars or more (9)
TenaamAuto	Ascription/ registration of the car of the main car user (OViN variable, in Dutch: HoofdAuto).	**Ownership company car**	
		No	Own name (1)
			Other person within household (2)
			Other person outside household (3)
			Unknown (5)
			Not asked; person: is below age 17, or has no licensure, or licensure is unknown, or is no main car user, or main car use is unknown (6-10)
		Yes	Company (4)
HHFiets	The number of non-e-bicycles in the household.	**Ownership household bicycle**	
		No	0
			Unknown (10)
		Yes	1..8
			9 bicycles or more (9)
Sted	Urban class residential municipality: average addresses density, measured per address within a 1 km radius.	**Urban density (addresses/km^2)**	
		High (≥ 1500)	Very high urban (≥ 2500) (1)
			High urban (1500-2500) (2)
		Medium (1000-1500)	Moderate urban (1000-1500) (3)
		Low (< 1000)	Low urban (500-1000) (4)
			Not urban (< 500) (5)
Jaar	Reporting year.	**Year**	

Note: Missing is transformed into a Missing value.

the research of OViN in 2013 and 2014, it was discovered by CBS that unclarity about the question led to a lower reported individual total bicycle ownership (OViN variable, in Dutch: OPFiets) for those years than previous years (about 10%). It resulted in a recommendation not to use the bicycle ownership data for those years (Statistics Netherlands, 2014). The reported ownership of a bicycle in the household was only about 1% lower, resulting in using the variable household bicycle ownership for our research, as this has likely a limited impact on our research.

The urban density is based on the average addresses density for the addresses in a residential municipality. This variable is transformed into a high, medium and low urban density category. A high urban density consists of a very high and high urbanity, a medium urban density of moderate urbanity, and a low urban density of the low and not urban class.

Inactive covariate (residential municipality)

Regarding residential municipalities (inactive covariate, Table 8), several municipal reorganisations have been made in 2010-2017. Table 11 shows the number of municipalities per year. Our operationalisation involves those municipalities (consisting of one or several cities) merged into an existing municipality, or several municipalities combined into a new municipality due to the reorganisation.

Accordingly, the municipality division of 2017 (388 municipalities) is used for the individuals in all the years to visualise the results about municipalities geographically. In the OViN data, the municipality names were slightly different spelt out in several years, which is firstly corrected. The residential municipalities are transformed based on the list of previous Dutch municipalities of 2010-2017. It should be noted that some municipality names in the list have slightly different spelling than in the OViN data, which is useful when using a code. The list is used based on the municipalities which stopped existing from 01-01-2010 until 01-01-2017 (column, in Dutch: Bestaan tot), showing which municipality (column Gemeente) has been merged into which municipality (column, in Dutch: Opgegaan in), which also shows in the value of that column if it is a new municipality (in Dutch: 'naamswijziging' in brackets). Sometimes, a municipality can be merged into a new municipality several times in this period. The latest merging is then used to correspond to the latest division. When several parts of a municipality are merged into several municipalities, individuals are all merged into the biggest municipality. This was the case for two municipalities, Boarnsterhim (191 individuals) into Leeuwarden and Maasdonk (64) into 's-Hertogenbosch. Only a few individuals (some of the mentioned ones) are thus placed in the wrong nearby municipality.

Table 11. Operationalised residential municipalities (OViN variable, in Dutch: Wogem) 2010-2017 (inactive covariate) on the individual level.

	Number per year								Transformed categories
	2010	2011	2012	2013	2014	2015	2016	2017	
Inactive covariate									
Residential municipality	431	418	415	408	403	393	390	388	388

4.6 Dataset processing

Based on the chosen sample and operationalisation choices of the travel behaviour, the data can be filtered and aggregated to obtain one record per individual, which is further explained in section 4.6.1, whereafter the final sample numbers per year are given in section 4.6.2.

4.6.1 Dataset explanation, filtering and aggregation

Before filtering and aggregating the data, it is essential to know that the OViN data comprises information for every stage, part of a trip (see section 4.4), in every row (one record). So, one individual can have multiple records when having more than one trip or in one trip multiple stages. Moreover, the matters of stages are recorded, and several options exist that do not match the travel behavioural definitions made in section 4.4. Furthermore, some individuals (and their characteristics) are not matching the defined sample or scope (section 4.3). The explanation of the dataset is further given in Appendix A.1. Besides, the explanation of the filtering (Appendix A.2) and aggregation (Appendix A.3) to obtain one record per individual with information about all stages is in a detailed way provided in Appendix A. This supplemental information about data processing is given as the OViN data is a comprehensive and extensive dataset, making it necessary knowing how to use this dataset for the reproducibility of research. This section briefly explains the filtering and aggregation steps which are done.

Individuals are first filtered out based on having values on certain variables which correspond to children (sample) or when individuals have not been away or only did non-daily travel (general aspects of travel behaviour) to match our chosen scopes.

For other individuals who need to be filtered out, existing or to-be-made (conditional) variables are used based on values at the stage level. This is because individuals can have (some) stages which do not correspond to our travel behavioural definition. When aggregating the individual's values on records of (conditional made) variables to obtain one record per person, they can be excluded once easily. So, the obtained values of (new conditional) variables after aggregation can be used for filtering or creating new measurements. The conditional variables are made to match the definition of general aspects of travel behaviour: partly non-daily travel (including foreign travel), and the travel behavioural mode set aspect (another mode set).

After aggregation, next to the stored aggregated (conditional) variables on the individual level, the other personal variables, like age, are stored on the stage level per individual in one record. Several measurements (travel behaviour and multimodality measurements) are created after aggregation based on the used and created variables, explained hereafter in section 4.7.

Dataset explanation, filtering and aggregation (detailed)

See Appendix A.

4.6.2 Final sample

After excluding cases by filtering before aggregation and filtering after aggregating on the stage level, measures can be created on the individual level (section 4.7). The final sample consists now of about 21,000-25,000 individuals per year (2010-2017), resulting in a final total sample of 183,618 individuals. It can be seen that a limited amount of individuals are excluded based on the second-mentioned (grouped) requirements in Table 12.

Table 12. Initial sample and final sample (after excluding cases) of individuals per year.

	Individuals per year								Total
	2010	2011	2012	2013	2014	2015	2016	2017	
Active covariate year									
Initial sample	44,165	42,338	43,307	42,350	42,600	37,350	37,229	38,127	327,466
After excluding for: children, has not been away, only non-daily travel (incl. foreign travel)									
	26,584	25,393	26,418	25,675	26,042	22,730	22,552	23,133	198,527
After excluding for (after aggregation): partly non-daily travel/foreign travel/other mode set									
Final sample	24,641	23,423	24,564	23,740	23,968	21,061	20,791	21,430	183,618

4.7 Measurements of (multimodal) travel behaviour

After the data operationalising and data filtering, during aggregation (section 4.6.1), the previously operationalised variables to be used for measuring travel behavioural variables (section 4.5) are aggregated on the individual level. So, via this, new measurements can be created for travel behaviour (indicators) and inactive covariates (multimodality measurements describing travel behaviour). The creation is thus done with the help of newly created variables or with existing variables (which are aggregated) into new variables based on the stage records per individual to define measurements. The created measurements are shown in Table 13 and are based on the earlier definitions given in Table 8. How the measurements are created is now explained, first for travel behaviour and then for the multimodality measurements.

The measurements for travel behaviour (indicators) are as follows computed during the aggregation step. The newly created dummy variables on the stage level called 'stage mode *name of mode* dummy' before aggregation got a value of one for that stage (of an induvial) if the particular mode (Rvm, see section 4.5; indicators) was used and a zero otherwise. The indicators, measuring the number of stages per mode, are created and calculated by summating the corresponding dummy variables for the records per induvial during aggregation. This resulted in the '# Walk stages', '# Bicycle stages', '# PT stages' and '# Car stages' variables for the model.

The multimodality measurement, which measures the number of modes an individual used (inactive covariate), is calculated after aggregating by creating a dummy variable, '*mode* used', for all the modes based on the defined indicator measurement variables. For instance, this dummy got a value of one for an individual when the '# Walk stages' variable, measuring the number of walk stages of an individual, is a one or higher, and zero otherwise. The number of modes an individual used is the sum of these dummy variables of all the modes.

The multimodality measurement OM_PI (inactive covariate), as theoretically and mathematically defined in section 2.4.5, uses the intensity of use 'number of stages per mode'. OM_PI is calculated after aggregation to the individual level. First, the variable 'total number of modes' is created and gets a value of 4 for every individual, which is the number of modes used in the mode set used in this research (N in (Equation 1)). In order to calculate the share (f_{ij} in (Equation 1)) of the number of stages (intensity of use) per mode (j), compared to the total stages made by that individual (i), the variable 'Sum stages' is created by summing the existing variables measuring the '# Walk stages', '# Bicycle stages', '# PT stages', and the '# Car stages' for that individual. Then, the 'Share *mode*' is calculated for every mode by dividing the number of stages per mode by the sum of stages. Then per mode, the formula from (Equation 1) is calculated and stored in a variable 'OM_PI *mode*' per individual. It is accounted for that if the share is 0, this variable gets a value of zero, as dividing to zero is impossible. Last, OM_PI is calculated by summating the 'OM_PI *mode*' values for all the modes of that individual.

Table 13. Created measurements (indicators and inactive covariates) on the individual level (using stages for travel behaviour).

Overarching variable group	Created measurement	Range
Indicators		
Travel behaviour (over one day)	# Walk stages	0 … max value
	# Bicycle stages	0 … max value
	# PT stages	0 … max value
	# Car stages	0 … max value
Inactive covariates		
Multimodality measurements of travel behaviour (over one day)	Number of modes	1 … 4
	OM_PI*	0 … 1

*Objective Mobility Personal Index. A higher value indicates a higher degree of multimodality.

4.8 Descriptives statistics of the final sample

This section is about the descriptive statistics of the final sample after data transformation (section 4.6) and measuring (section 4.7) based on the conceptual model (section 4.2). The statistics are mainly described by using Table 14. All statistics are based on all the years and for all the residential municipalities. The distribution over the years is also shown (see inactive covariate year), where all years are almost equally represented. Moreover, the number of municipalities (inactive covariate residential municipality) is shown. Some variables (covariates) have missing values (ethnicity, education level, occupation, household income, licensure), but their amount per variable is relatively low. First, the indicators and covariates are explained, potentially using histograms to show the frequency distributions. After these descriptions, the representativeness of the sample is described. Moreover, the correlations between those variables are examined, as explained earlier in the conceptual model. Last, the multimodality measurements are explained, also with the use of histograms.

The indicators and covariates are subsequently described for the sample (Table 14), consisting of 183,618 individuals who made 681,025 stages.

For the travel behaviour (indicators), it can be seen that there are, on average, about twice as many car stages (1.8) over one day made by individuals than walking (0.8) or bicycle (0.9) stages. Moreover, PT stages have the lowest prevalence (0.2 stages over one day) and the lowest standard deviation of 0.8. The other modes have a standard deviation of around 1.6.

The sample (covariates) is characterised as the following. Gender, part of socio-demographics, is almost equally represented. For the sake of simplicity, the continuous covariate variable age (and the other continuous variables) only shows the mean and the standard deviation. The mean age is 49 years, and the frequency distribution (Appendix B) shows a nominal distribution. Native (non-immigrant) citizens account for 85% of the sample. Moreover, many individuals have finished pre-vocational education or higher (93.3%). The 'other' household composition category is characterised by 47% of the respondents, where other household compositions are mainly compositions with three persons or more, compared to one-person and couple households which account together for 54% of the sample. Concerning occupation, there are many employed (57%), and many who have another occupation than listed (34%), which are individuals who are working in their own household, are unemployed, are unable to do work or are retired, compared to students (6%), and some missing values (3%). Household incomes from €10,000 - €40,000 account for 84% of the sample. About the mobility resource variables, many (≥ 88%) individuals of the sample own at least one of the following: licensure, a household car or a household bicycle. Concerning company cars, about 6% have a company car. Regarding the built environment, many live in a place with a high address density within a one km radius and a low address density, respectively 45% and 36% of the respondents.

Table 14. Descriptive statistics of the sample (N = 183,618; M = 681,025 stages).

Variable		
Indicators		
Travel behaviour		
# Walk stages (over one day)	Mean (SD)	0.8 (1.4)
# Bicycle stages (over one day)	Mean (SD)	0.9 (1.5)
# PT stages (over one day)	Mean (SD)	0.2 (0.8)
# Car stages (over one day)	Mean (SD)	1.8 (1.8)
Total stages	Mean (SD)	3.7 (2.2)
Active covariates		
Socio-demographic variables		
Gender (%)	Male	47
	Female	53
Age (18+)	Mean (SD)	49 (17)
Ethnicity (%)	Native	85
	Immigrant	15
	Missing	~0
Education level (%)	No education	0.7
	Primary education	4.3
	Pre-vocational education	21.5
	Vocational & higher secondary education	38.0
	Higher vocational & university education	33.8
	Missing	1.7
Household composition (%)	One-person household	18
	Couple household	36
	Other household	47
Occupation (%)	Employed	57
	Student	6
	Other occupation	34
	Missing	3
Household income (%) (standardised)	< €10,000	3
	€10,000 - €20,000	26
	€20,000 - €30,000	38
	€30,000 - €40,000	20
	€40,000 - €50,000	7
	≥ €50,000	5
	Missing	~0

Table 14. (continued). Descriptive statistics of the sample (N = 183,618; M = 681,025 stages).

Variable		
Mobility resource variables		
Licensure (%)	No	12
	Yes	88
	Missing	~0
Ownership household car (%)	No	11
	Yes	89
Ownership company car (%)	No	94
	Yes	6
Ownership household bicycle (%)	No	5
	Yes	95
Built-environment variables		
Urban density (%)	High (≥ 1500 addresses/km^2)	45
	Medium (1000-1500 addresses/km^2)	19
	Low (< 1000 addresses/km^2)	36
Time		
Year (%)	2010	13
	2011	13
	2012	13
	2013	13
	2014	13
	2015	12
	2016	11
	2017	12
Inactive covariates		
Multimodality measurements of travel behaviour		
Number of modes (over one day)	Mean (SD)	1.4 (0.6)
OM_PI (over one day)	Mean (SD)	0.2 (0.3)
Built-environment variables		
Residential municipality	N	388

Note: some column values may not add up to 100% or show 0% due to rounding.

Besides the shown statistics, some ordinal variables (education level, household income, urban density) are visually displayed in Appendix B. It can be seen that all are about normally distributed, meaning that they can likely be regarded as continuous (ratio) variables, especially the latter two variables, where the categories represent individual numeric values but are measured in categories. However, for urban density, the nominal distribution is less clear because the five original categories are likely more nominal distributed, but we used them after merging into only three categories. As the sample is high, a less clear nominal distribution has limited implications when regarding it as nominal in the upcoming model (see section 4.9).

Next, as seen in Appendix B (representative analysis), our sample is compared to the descriptives of the entire Dutch population. It should be noted that not all values are comparable to ours due to other definitions or operationalisations of Statistics Netherlands. So, no independent t-test is done, a statistical test for determining whether there is a significant difference between the means of the two groups[2], as this would give biased information. However, when comparing the distribution values for the variables and within the categories, the sample is likely highly representative for gender, age, household income, licensure, ownership of household car, ownership of company car, and urban density; quite representative for ethnicity and education level, somewhat comparable for household composition and the employed occupation category (these have comparability issues); and no information could be obtained for the other occupation categories and ownership of a household bicycle.

As part of the descriptive statistics, it is also checked if the active covariates do not overlap too much, as their individual effects are less captured when overlapping too much (explained in section 4.2). Because some variables are categorical, dummy variables are created. For instance, regarding gender, a variable gender_male obtains a one if the gender is male and zero otherwise, and gender_female the other way around. The continuous variables and assumed continuous variables are not dummy coded. Age is shown continuously and nominal (per year). The correlations can be seen in Appendix B. The correlations between several variable categories are really low (including urban density and year(s)). The others are now explained.

Generally, socio-demographic variables show very low correlations between each other as they measure all different phenomena, in our case, mostly (~ < 0.3). For age, a moderate correlation (0.63) is seen with the other occupation category, meaning that someone older is more likely to have another occupation than listed and the other way around. Age has some more low correlations (between 0.3 and 0.5) with household composition and occupation categories. Still, age is regarded as a different phenomenon than those other variables.

On the other hand, it might be that some variables capture the same for the mobility resource variables, as it is about owning mobility resources. The highest correlation in this set is about 0.5, and just before the boundary of being very low and moderately correlated. This correlation is between having a licensure and owning a household car. So, when owning a licensure, someone is likelier to own a household car and vice versa (no causality can be assumed). However, it is intuitive that people own a household car only (mostly) when having a car licensure. Both variables are still included, as the correlation is very low.

The distributions of the multimodality measurements (inactive covariates) are shown in Figure 9 and Figure 10, besides the statistics in Table 14.

First, the distribution of the number of modes is explained. Figure 9 shows that most respondents (~123,000; 67%) used one mode over one day, and for two, three and four modes, fewer and fewer respondents have used that amount. This is also reflected in the mean value of 1.4, but the standard deviation is 0.6 (Table 14). About 49,000 respondents (27%) used two modes over one day, about 11,000 respondents (6%) used three modes over one day, and the fewest amount of respondents (~900; 1%) used four modes over one day.

Second, the measurement OM_PI distribution (Figure 10) shows that the same amount of people who used one mode over one day have an OM_PI of 0.00 (67%) because one mode is used for all stages, which is the least multimodal. The average value is 0.2, with a low standard deviation of 0.3 (Table 14). Hardly anyone has a value of 1.00 (< 1% of the respondents), which means that four modes

[2] https://libguides.library.kent.edu/spss/independentttest

are used, and all modes are used with equal intensity. About 17,000 (9%) of the respondents have a value of 0.5, so two modes are used in a balanced way. Values around these could mean that, for instance, two modes are somewhat used in a balanced way, and if a third mode is used, it is imbalanced used. About the same amount of respondents have a value of 0.46 (~16.000). About 2,300 respondents (1%) have a value of 0.75, which is when three modes are used balanced. Values around 0.75 reflect individuals who probably used three modes but just not in a balanced way. It also shows a small (local) peak, representing few respondents.

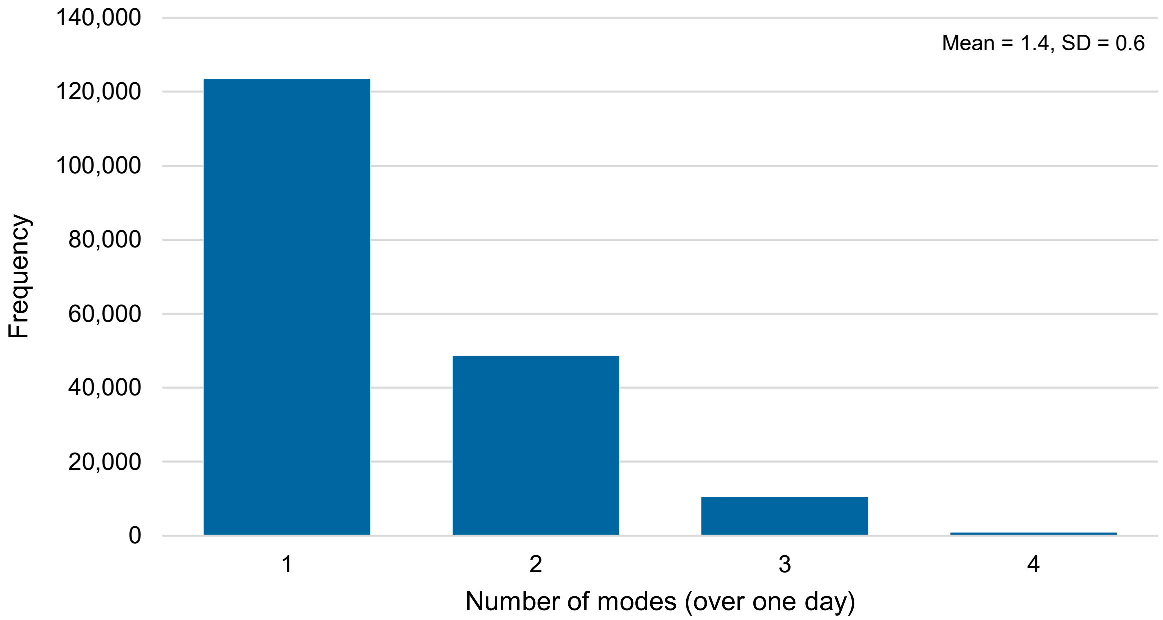


Figure 9. Frequency distribution number of modes used on the stage level by individuals (N = 183,618).

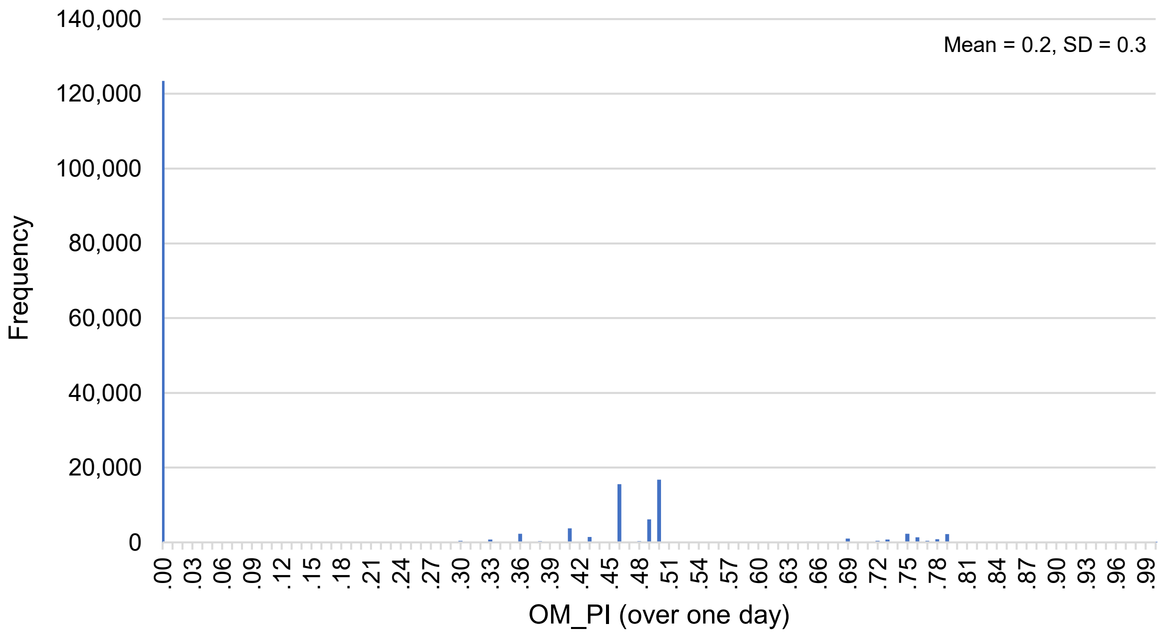


Figure 10. Frequency distribution OM_PI on the stage level of individuals (N = 183,618).

4.9 Modelling strategy

After all the previous, the model can be specified and estimated using the final data. The process is described in the modelling strategy. The modelling strategy includes describing the model specification and estimation (section 4.9.1) and the final model selection (section 4.9.2).

4.9.1. Model specification and estimation

Before it is explained how the model is specified, the estimation, including a robustness check, is first explained. After those aspects are elaborated, it is explained how to check if the estimation process has succeeded.

The measurement model with indicators only is estimated using various latent classes (1-10) from which the model fits, and the to-be-explained criteria (section 4.9.2) will be compared per model with several latent classes to choose one model solution. Afterwards, the covariates (structural model) are added and estimated for the chosen model with a number of classes to obtain the final model. See section 4.2 for the description of these explained used terms from the conceptual model.

The re-estimation is a common approach (e.g., Molin et al., 2016; Weller et al., 2020). Generally, running only a measurement model takes less time, making it more convenient to run several solutions before choosing one. An appropriate aspect of this process is that the measurement part of the model, where the chosen solution is based on, hardly changes when including the active covariates in this 'active covariates method' in general (Vermunt & Magidson, 2016). However, selecting the classes based on a model that is not used later (it is re-estimated) may lead to flawed results due to potentially changing measurement models (Vermunt, 2010; Weller et al., 2020). Based on the potentially flawed outcome, a robustness check is done to see if the measurement model is fairly the same in the final model, including the structural model, compared to the measurement model only, which is used to choose the number of classes. The robustness check is, however, a subjective task, as no guidelines in magnitudes exist. Nevertheless, doing this check gives some insight into how robust (stable) or sensitive (changing) the outcome is due to adding covariates.

The model with indicators only and the model with indicators and covariates are estimated using Latent Gold 5.1 (2016); see also Vermunt & Magidson (2016). Some specifications from default are specified to the following.

Firstly, complete information maximum likelihood estimation is used. In Table 14, no missing values are present for the indicators, but a limited amount of missing descriptive values are present at some covariate values. Latent Gold handles the missing values by handling the effects of this individual on this covariate as zero (nominal variable), or it imputes the sample mean (numeric) for that covariate variable. In this way, the individual, which is placed in a class based on their indicators values, values can also be used to describe the outcomes with the other covariates which have values.

Second, the technical estimation settings are changed. Because of the possibility of ending up with a solution, which is a local optimum, both the measurement models and the final model, including the measurement and structural model, are estimated with 160 sets of random starting values, where 250 iterations are specified for each. The iteration limits for EM (Expectation-maximization) and N-R (Newton-Raphson) algorithms are the same as the default settings; see Vermunt & Magidson (2016) for more details on this.

Third, the scale types are specified. The indicators (stage rates) are set to count in the measurement model. After the final-class solution is chosen based on the measurement models, the covariates are added to the measurement model to re-estimate it. The active covariates are specified as nominal (including year) and numeric (continuous variable age), and some nominal variables are also set as numeric to limit the number of parameters needed in the model (education level, household income, urban density). Concerning year, as a limited amount of time points are present, and no interaction effects are included, it can be included as nominal (instead of being regarded as ordinal or continuous). The latter-mentioned variables, set to numeric, are handled this way as they are regarded as normally distributed (section 4.8). Regarding the inactive covariates, the multimodality measurements are numeric, and the residential municipality variable is nominal.

After running the models, all the specified and estimated models are checked for convergence, meaning an optimal solution has been found. As explained, after the measurement models are run, the chosen class solution is re-estimated for the measurement model by including the structural model with covariates to obtain the final model. Moreover, the latter chosen class solution (measurement model with indicators) and the final model with added covariates (structural model) are both run twice to check if the previous outcome, regardless of using random start sets (see before), was not a local solution by checking if the parameters and Log-likelihood, have about the same values. If not, for both, the first runs will be used; otherwise, it will be reported and again double-checked with other random start values.

4.9.2. Final model selection

The determination of the final latent class model, with the final number of classes, is chosen based on the criteria for determining the final model. The criteria are first explained, and second, the determination is explained.

Criteria for determining the final model

When comparing models with different classes (solutions), several criteria are used to determine the number of classes. In the ideal situation, no criteria should solely be used to decide. So, these criteria are used jointly, but the choices of using which one can be dependent on the model outcomes and their resulting values on criteria. There is also room for interpretation when choosing the number of classes in general, next to the room for interpretability for the criteria outcomes. In short, before diving deeper into them, the model fit criteria are primarily for choosing, and the criteria based upon these model fit criteria are more regarded as secondary. Diagnostic criteria can guide the choice and are important for consideration, so these are named tertiary criteria. The interpretability of results is useful for consideration among all criteria.

Below, the used decision criteria are depicted, argued for, and overall explained:

Used decision criteria (jointly use) to decide upon the number of classes

		(Global) model fit criteria
Primary	1.	Likelihood ratio chi-squared statistic, L^2: low(est) value among models is generally preferred
	2.	Bayesian information criterion, BIC(LL): low(est) value among models is generally preferred
Secondary	3.	% Reduction in L^2 compared to 1-class model: high difference with the previous class model value on this criterion is generally preferred
		Diagnostic criteria
Tertiary	4.	Smallest class size: at least 5% is generally preferred (in this case)
Among all	5.	Interpretability of results: number of classes should be generally intuitive

Before explaining the criteria (primarily up to tertiary) one on one, some general information about these criteria and why only global fit measures are used is explained. The per cent reduction in L^2 compared to the 1-class model (3) is a less formal approach that can complement the more statistically precise L^2 (1) and BIC (2) approaches, as explained below after this paragraph. More substantive criteria can also be used (Weller et al., 2020), like the smallest class size (4) as diagnostic criteria, next to the criteria primarily for selecting the solution based on model fit.

Next to global model fit measures (1-3), and diagnostic criteria for guiding the choice (4), the local model fit is sometimes used to determine the number of classes. Especially when the global model fits cannot differentiate well enough to find a suitable solution, bivariate residuals (BVRs) can be used (Molin et al., 2016). These are estimates for an improvement in model fit (L^2) when direct effects between indicators are added, which is a relaxation of the local independence assumption. Values higher than 3.84 would indicate a significant covariation between a pair of indicators, and having no (or a few) significant associations left is preferred. The sum of the BVRs can thus indicate the amount of association left between the indicators when accounting for the latent variable (Kroesen, 2019). Lower values are thus generally preferred for this criterion. Irrespective of using which measure for discriminating on the BVRs, when using models with many cases, having no significant associations left in the BVR estimates is challenging. This method will not be used in this case, especially when criteria 1-4 are sufficient in choosing the number of classes. So, it is not mentioned as an additional criterion used in the depicted criteria above.

The four criteria used (see depicted above) are now explained, and the fifth criterion, interpretability, is discussed for all of them.

First, as explained by Magidson & Vermunt (2004), the most widely used statistical criteria to determine the number of classes is the likelihood ratio chi-squared statistic (L^2). It assesses the fit of the latent class models by looking at the extent to which the maximum likelihood (ML) estimates for the expected frequencies of values differ from the observed frequencies. A value of L^2 is sufficiently low for the model to be attributable to chance, and the lowest value between model solutions is generally preferred because a value of L^2 deviating from zero (perfect model fit) shows the amount of association that remains unexplained by the model.

Second, another approach for assessing the model fit is one of the information criteria, the Bayesian information criterion BIC is to compare models (Weller et al., 2020). BIC is one of the most used among researchers, according to them. Nylund et al. (2007) also explain that some researchers consider it the most reliable model fit indicator. Because, per definition, a model with more classes has a better fit, a trade-off with parsimony is needed. The BIC assesses both model fit (Log-likelihood) and parsimony, which is the number of parameters needed to estimate the model. Per definition, a bigger model needs more parameters (Magidson & Vermunt, 2004). A model with a lower BIC value is generally preferred over a higher one.

Third, another approach is also used to determine the number of classes. This approach uses the L^2 of the 1-class model (baseline) and compares it with the L^2 of a higher-class solution (Magidson & Vermunt, 2004). A reduction means a higher model fit of that class solution compared to the baseline. The reduction per cent measures the association explained by the model with a certain number of classes because the 1-class model mostly has inadequate model fit and shows the amount of association in the data. As De Haas et al. (2018) explained, when comparing the reduction values of several model solutions, if the difference between the value of this criterion of a class solution and the value of this creation for the previous model solution with one class less becomes relatively small, it is no longer necessary and justified to add an extra class to the model.

Fourth, the smallest class size, a diagnostic criterion, is next to the criteria primarily for selecting the solution used. Ton et al. (2020) also used this and advocated that classes above 8% are preferred.

Weller et al. (2020) explain that 5% might be suitable. As a large dataset with many cases is used, 5% as a guideline is chosen in this case. Otherwise, it can be argued that the style is more of a niche style, and more classes hamper the interpretability of more general styles (Sinha et al., 2021).

Determination of the final model

The rationale for choosing the number of classes is given after estimating the measurement model, as explained before, using the criteria just explained. The models (with 1-10 classes), which all have been converged ($p < 0.00$), are shown. Several criteria are listed in Table 15.

Based on the L^2 and BIC(LL), the criteria keep decreasing. It means more than ten classes should be needed to find an optimal value. However, when more classes are used, there is also a risk of overfitting the data, and it is less able to generalise the results as the model becomes too specific. As interpretation and communication of the classes are also important, the percentage of reduction in L^2 compared to the 1-class model is seen as another good measure. From the 5-class model onwards, the difference between the values of this criterion between the n-class solution and the n+1-class solution is relatively low (3% or lower). So, the latter models almost explain the same amount of association in the data, and the improvement of model fit stagnates. In Figure 11, the criterion % Reduction in L^2 compared to 1-class is visualised, and the line from the 6-class models onwards becomes almost linear and relatively flat. The difference between the criterion value between the 4-class and 5-class models is quite substantive. Moreover, the 5-class model has the smallest class size of 10%, which is higher than the set minimum of 5%.

After all, although the 6-class model has an appropriate smallest class size of 7.2%, the low gained improvement in model fit (L^2) compared to the 5-class model is not preferable. Because the 5-class model has better parsimony (fewer parameters needed) and on the previous set reasoning, the 5-class model is selected as optimum (shown in boldface in Table 15 and by the circle in Figure 11).

The final model is re-estimated by including the structural model with covariates in the 5-class solution. The measurement model is highly robust in this 5-class solution when covariates are added, as there are almost no changes in class size and the values on the indicators, or limited changes in the values (lower than 0.1; not shown). The measurement model outcome (from the model with indicators and covariates), with the magnitude of the values, is given in Chapter 5.

Table 15. Model fit evaluation from the models with indicators only (measurement model) (N = 183,618).

	Criteria:	Model fit*				Diagnostic
Number of classes	**Npar**	**LL**	**BIC(LL)**	**L^2**	**% Reduction in L^2 comp. to 1-class**	**Smallest class size (%)**
1	4	-1,022,347	2,044,742	643,545	0	100.0
2	9	-893,171	1,786,451	385,194	40	44.4
3	14	-833,036	1,666,241	264,923	59	19.8
4	19	-812,119	1,624,469	223,090	65	10.5
5	**24**	**-794,177**	**1,588,645**	**187,206**	**71**	**10.0**
6	29	-785,348	1,571,047	169,547	74	7.2
7	34	-779,936	1,560,284	158,723	75	4.8
8	39	-774,924	1,550,321	148,700	77	4.2
9	44	-772,020	1,544,573	142,891	78	2.9
10	49	-769,705	1,540,004	138,262	79	2.3

Abbrev.: Npar: Number of parameters, LL: Log-likelihood, BIC(LL): Bayesian Information Criterion (based on LL and Npar), L^2: Likelihood ratio chi-squared statistic.
* Higher LL, lower BIC or L^2, and higher % Reduction in L^2 comp. to 1-class represents a better model fit.
Bold = Chosen solution.

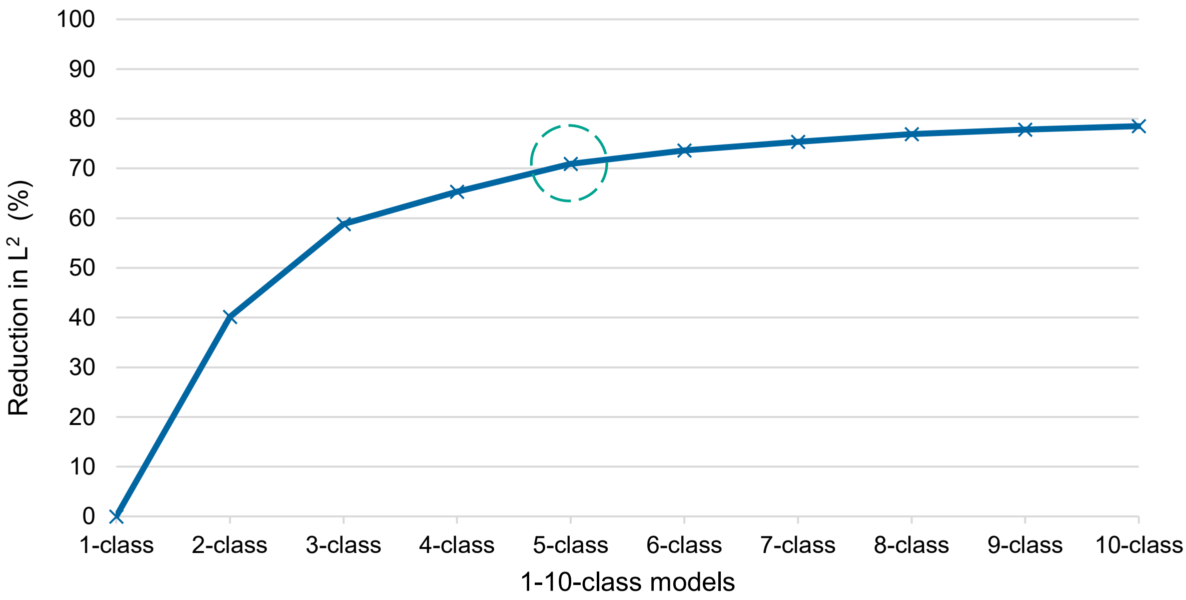


Figure 11. % Reduction in L^2 compared to 1-class model (circle = chosen class solution).

5. Results

This chapter comprises the results after using the previous chapter's data sample and model estimation. First, the trends of the main used variables, namely travel behaviour determinants and travel behavioural variables, are overviewed for several years (section 5.1). Second, the captured multimodal travel patterns of the identified travel user classes are explained and how multimodal they can be characterised (section 5.2). Third, the effects of determinants on being likely a member as a travel user of those classes with specific modal travel patterns are elaborated upon (section 5.3). Fourth, the spatial distribution across municipalities of the class membership of travel users is visualised (section 5.4). Fifth, the development of the class sizes and classes' multimodal travel patterns are investigated (section 5.5).

5.1 Trend overview of travel behaviour determinants and travel behaviour

This section describes the trends of the active covariate values (determinants) and travel behaviour-related variables used as indicators or inactive covariates (see section 4.2) in the upcoming shown LCA model to embed the development of modal travel patterns more into context.

The overall descriptives of (potential) determinants are shown in Table 14 in Chapter 4. Their average values (also for the nominal variables, which provide less accurate results if the mean is used) and standard deviations hardly change over the years (not shown). It is assumed that the distribution within nominal variables also did not substantially change. As it is earlier assumed that the sample is quite representative of the target population (Dutch residents), this is likely also the case for all the years separately, as in a few years, limited major shifts in demographics or other aspects can occur. Although, if overall averages or distributions are similar, it should be noted that shifts could have occurred within the distribution, and the same aggregate values remained.

The average trends of travel behaviour-related variables (over one day), as the total stages, and multimodal measurements, as the number of modes used and OM_PI, are shown in Figure 12 and Figure 13. The average total stages over one day slightly declined from 2010-2017 (about 3.9 to 3.6); see Figure 12 (a). Although a trend depends on how it is framed, as these lines might look different when more years would have been depicted, it can be seen as relatively stable, as the range of values is relatively low with about 0.3 stages. The average number of modes per year was also stable in 2010-2017, as every year has a value of about 1.4 (Figure 12, b). OM_PI depends on the number of modes used and the intensity of stages per mode (section 4.7). Like the average total stages, the average stages per mode per year were also highly stable from 2010-2017 (not shown). So, OM_PI is also consistent in this period, with an average value of 0.17 (Figure 13).

Based on the previously shown consistency of travel behaviour over the years and that the sample every year is comparable and assumed to be representative of every year's target population, it can be considered that the travel behaviour patterns are highly comparable (and representative) over the years. Based on this observation, it is interesting to see, unless average comparable descriptives, which 'hidden' (multi)modal travel user groups exist and if the classes have developed over time, which will be investigated in the upcoming sections.

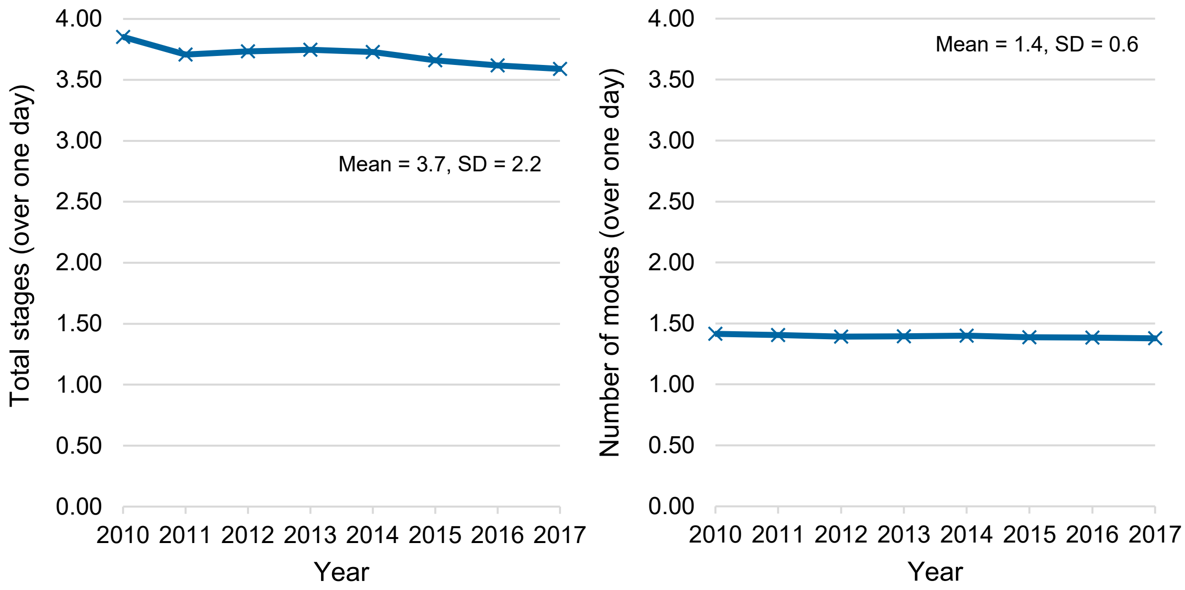


Figure 12. Trend overview 2010-2017 of average total stages (a, left) and average number of modes (b, right).

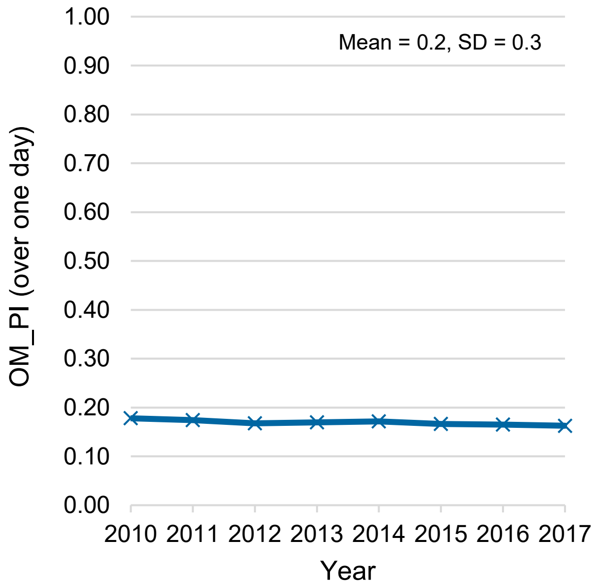


Figure 13. Trend overview 2010-2017 of OM_PI.

5.2 Latent classes: profiles of multimodal travel patterns and degrees of multimodality

The final model is estimated using the final data sample and specification and estimation process described in the previous chapter, based on the conceptual model in section 4.2. Some trends of input variables are already explained in section 5.1. This section shows the latent class profiles ('Profile output'), consisting of multimodal travel patterns and degrees of multimodality, to give names to the identified classes hereafter. The 'Profile output' of the final model with indicators and covariates is shown in Table 16 to show the within-class distribution of variables/values of the emerged latent classes. The profiles are only shown for the indicators and some inactive covariates to define the travel patterns of the classes based on the multimodal modal usages. Appendix C; latent class profiles, shows the distribution for other model variables. The mode usage (indicators) is thus shown per class, and the values in the table represent the mean value of the indicator per cluster. The inactive covariates are shown to define how multimodal classes are, based on the multimodal measurements of the people in classes. Moreover, Figure 14 shows the distribution of values or range of values per class for the multimodality measurements. Besides, the class sizes in the table reflect the mean class size across all the years, and the sample distribution of the used variables is shown in the table to interpret values in the classes easier. The average number of stages travelled per class is also calculated and shown in the table. Among the classes, the values in boldface indicate the highest values across the classes. After identifying the patterns, these can be used in section 5.3 to investigate the influence of characteristics on being a member of the defined classes.

Moreover, the Wald statistic is shown (based on the model parameters) to assess if a parameter is significant from zero (Appendix C; latent class model parameters) and thus has a significant relationship with the latent class variable. In this case, the latent class variable is a 5-part categorical variable, as five classes are used in this model. Based on the Wald statistic, using the 95% confidence interval ($p < 0.05$), it can be seen that all indicators are significant. So, all indicators are highly significant and thus different between the classes.

Five classes are identified based on Table 16 and are named based on the indicators and multimodality measurements (inactive covariates). To better define multimodality, the distribution of the multimodality measurements value (or ranges) is visualised in Figure 14, next to the given mean values in the table. The identified classes are now explained.

1. Car exclusive (C)

The first class (1), Car exclusive (C), is defined by the highest share of people (41%). On average, people take 3.0 stages by car over one day (the same as the total stages over one day), the highest among all classes and about twice as much as the sample mean (1.8 stages). But the total amount of stages over one day (3.0) is close to the sample mean of total stages (3.7 stages).

This class is also characterised by the lowest multimodality measurements with, on average, 1.0 used modes, which reflects the on average unimodal car exclusive behaviour, also seen in the within-class distribution of the number of modes in the figure. Figure 14 (a) shows that almost 100% of individuals in this class used one mode. The other multimodality measurement, OM_PI, has an average value of 0.02 in the table, and almost all individuals in this class have a value of zero (Figure 14, b). These measurements are quite close to the sample mean values of 1.4 and 0.2, respectively. So, when being in class (1), it is more likely to have many car stages, and low multimodality measurements characterise the travel behaviour over one day.

2. Bicycle mostly (B)

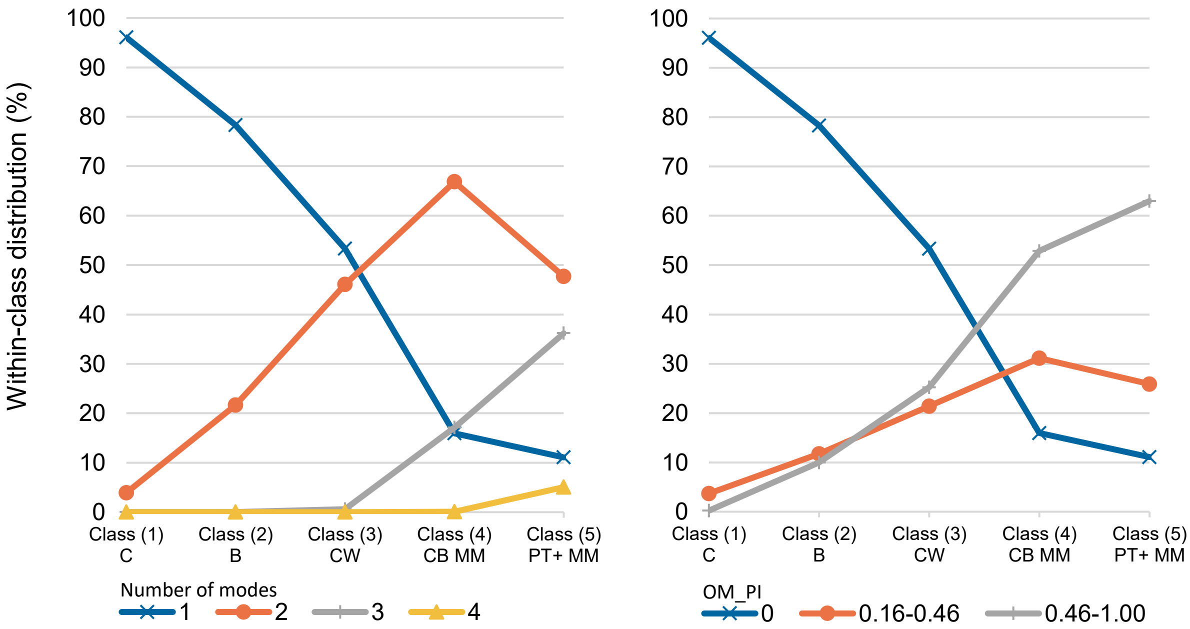
The second class (2), Bicycle mostly (B), is characterised by relatively many people (20%) and by the highest number of bicycle stages among the classes, with an average of 2.7 stages over one day, which is thrice as much as the sample mean value. The walking trips are relatively low, with 0.4 stages over one day on average (twice less than the sample mean), so this class is called Bicycle mostly. On average, 3.1 stages are made over one day, comparable to the sample mean of total stages.

The multimodality measurements are comparable to the sample mean values. This class has, on average, 1.2 modes used, reflecting the more unimodal behaviour, and an OM_PI of 0.10. Figure 14 (a) shows that most people use one mode (about 80%), and some used two modes in this class over one day. This is reflected in many people having an OM_PI of zero (about 80%), or OM_PI belonging in the range of 0.16-0.46, and some in the group of 0.46-1.00 (the latter two about 10%), see Figure 14 (b).

Table 16. Profiles 5-class solution: within-class distribution/values from the latent class model with indicators and covariates; indicators and inactive covariates of multimodality measurements are only shown.

	Class[a]	(1) C	(2) B	(3) CW	(4) CB MM	(5) PT+ MM	Total sample [b]
Class size (%) N = 183,618		**41**	20	17	12	10	100
Indicators							
Travel behaviour (over one day)							
# Walk stages (Wald = 54235, p < 0.00)	Mean	0.0	0.4	**2.2**	0.4	**2.8**	0.8
# Bicycle stages (Wald = 12041, p < 0.00)	Mean	0.0	**2.7**	0.0	**2.0**	1.0	0.9
# PT stages (Wald = 8445, p < 0.00)	Mean	0.0	0.0	0.0	0.0	**2.2**	0.2
# Car stages (Wald = 29940, p < 0.00)	Mean	**3.0**	0.0	1.4	**2.3**	0.5	1.8
*Total # stages**	*Sum of means*	*3.0*	*3.1*	*3.6*	*4.7*	*6.5*	*3.7*
Inactive covariates							
Multimodality measurements of travel behaviour (over one day)							
Number of modes	Mean	1.0	1.2	1.5	2.0	**2.4**	1.4
OM_PI	Mean	0.02	0.10	0.22	**0.44**	**0.54**	0.2

a: (1) C: Car exclusive, (2) B: Bicycle mostly, (3) CW: Car + Walk, (4) CB MM: Car + Bicycle Multimodal, (5) PT+ MM: Public transport+ Multimodal.
b: Data are pooled for 2010-2017.
* Calculated on own.
Bold = The highest sizes/means for a variable compared to other classes.



(1) C: Car exclusive, (2) B: Bicycle mostly, (3) CW: Car + Walk, (4) CB MM: Car + Bicycle Multimodal, (5) PT+ MM: Public transport+ Multimodal.

Figure 14. Profiles 5-class solution: within-class distribution of number of modes (a, left) and OM_PI (b, right); inactive covariates are shown.

3. Car + Walk (CW)

The third class (3), Car + Walk (CW), is characterised by 17% of the people. It has a relatively high amount of walk stages (2.2 over one day on average) among the classes compared to the sample mean (1.8 stages). Moreover, this class has 1.4 car stages over one day on average, which is somewhat 'in the middle' compared to other classes, but about the same as the sample mean (1.8 stages). Besides, this class is even more comparable to the sample mean, showing that about 3.6 stages on average in total are made over one day, compared to the sample mean of 3.7. This class is labelled CW because it is likely that the car is used as the primary mode, and car stages are complemented by walking on other trips.

Moreover, this class is characterised by multimodality measurements, which are just about the sample mean values. It has an average of 1.5 modes used in this class, comparable to the two potentially used modes in this CW class (bimodal behaviour). Figure 14 (a) shows indeed that the number of individuals using one or two modes is almost comparable in shares (about 50%). The OM_PI is, on average, 0.22. The amount of individuals having an OM_PI of zero (one mode used) is lower compared to the previous two classes, about 55% (Figure 14, b). The two higher ranges of values of OM_PI measurements are also again quite comparable in size (about a share of 22%), and these values are thus somewhat more represented in this class on average than in the previous two classes.

4. Car + Bicycle Multimodal (CB MM)

The fourth class (4), Car + Bicycle Multimodal (CB MM), with somewhat fewer persons (12%) than the (3) CW class, shows a high amount of car stages (2.3 over one day) but also a high amount of bicycle stages (2.0 over one day on average) compared to other classes and the sample mean values (1.8 and 0.9 stages respectively). These stages are also somewhat complemented by walking (0.4 stages on average over one day), twice as low as the sample mean. The average total stages over one day is 4.7, somewhat higher than the mean value of 3.7. This class is labelled multimodal (MM), although two primary modes are used (C + B), and PT is not used on average. While PT is commonly characterised by multimodal behaviour, this class is now somewhat complemented by walking as a 'third' mode, showing the mode diversity in the case if three modes are used.

The second-highest average multimodality measurements occur in this class. The average number of modes (2.0) and OM_PI (0.44) over one day is considerably higher than the sample mean values of 1.4 and 0.2, respectively. The number of modes, on average (2.0), reflects that many used two modes out of the three maximum modes over one day used in this class. Figure 14 (a) also shows that most individuals used two modes in this class (about 70%), and about equal shares used one or three modes (about 15%). In Figure 14 (b), the highest OM_PI range of 0.46-1.00 is now most represented in this class; somewhat more than 50% of the individuals have a value in this range.

5. PT+ Multimodal (PT+ MM)

The fifth class (5), Public transport+ Multimodal (PT+ MM), has the most diverse mode usage but the lowest class size (10%). There are, on average, 2.2 PT stages over one day, which is the highest among all classes and considerably higher than the sample mean value of 0.2 stages, so this class is called PT. The other used modes are named in the +. Also, this class has the highest amount of walk stages (2.8 over one day on average), with a mean sample value of 0.8 stages. Bicycle stages complement these two modes, with 1.0 stages over one day on average and comparable to the sample mean, and it is complemented by car stages (0.5 over one day on average), which is relatively lower than the sample mean of 1.8. The highest average of total stages over one day of 6.5 characterises this class. This class is also labelled multimodal (MM) as it has the highest multimodality measurements.

On average, 2.4 modes are used over one day (from the four modes with a trip stage value), and OM_PI is 0.54. Figure 14 (a) shows that most people used three (about 50%) or two (about 35%) modes, and very few people used one (not multimodal behaviour, about 10%) or four (very multimodal behaviour, about 5%) modes. Like the previous class (4) CB MM, the highest range of OM_PI (0.46-1.00) is most represented in this class, now somewhat more represented with about 65% of individuals in this class, see Figure 14 (b).

5.3 Influence of determinants on latent class membership

The classes and their travel patterns are already defined (section 5.2). Now, the influences of determinants on latent class membership are explained. Table 17 shows the 'ProbMeans' output (also in Appendix C; latent class membership). Moreover, the class sizes are given again, which reflect the mean class size across all the years. This ProbMeans output shows the distribution of the latent class membership probabilities given the covariate value while holding the other covariates at their mean value. So, the outside class distribution can be seen (sum over rows is 100%), where an individual with a particular covariate category (e.g., being a male or female) has a higher chance of belonging to a travel user class with a higher probability. Also, the distribution for the missing values of covariates is given. However, as the sample shares of missing values are minimal, these effects are not interpreted. The inactive covariate residential municipality and the covariate year are explained later (sections 5.4 and 5.5, respectively).

It can be seen that all active covariates are also significant and thus affect class membership significantly, based on the explained Wald statistic (section 5.2). So, these active covariates can be assumed to be exogenous and used to predict class membership with specific travel patterns. Moreover, the size of the Wald statistic shows the strength of the effect of a covariate. A higher Wald statistic indicates that the founded relationship is relatively strong. The model parameters show (Appendix C; latent class model parameters) that the strength of the variables ownership of a household car, ownership of a household bicycle, and ownership of a company car is (relatively) strong (Wald = 5225, 2569, 1600, $p < 0.00$). Remember that these variables cannot be assumed entirely exogenous (section 4.2), and the effect is thus likely somewhat less. Moreover, the effect of occupation on latent class membership is quite strong (Wald = 3892, $p < 0.00$), and there is a somewhat strong effect of the urban density variable (Wald = 1953, $p < 0.00$).

Based on the showed results in Table 17, the following notions are important.

It should be noted that many covariate categories show overall a higher probability of belonging to the first class than to the second class, and the second than the third, et cetera because the first class has the highest class size, which decreases per class (2 to 5). When a covariate category deviates from decreasing probabilities for membership of classes with decreasing class size, this could point towards an interesting relationship. So, when it is said that probability values differ based on the 'overall seen trend' of class size, the previous reasoning is meant. Nevertheless, the size of the probability values is not one on one comparable to the class sizes, nor does it point towards a comparable distribution of values.

An example explains another notion of the effects of covariate categories on class membership. Given that an individual is a male, it has a probability (prob.) of 47% of being in the (1) C class while keeping the other covariate values on their mean level. In other words, of all males, 47% are in the (1) C class. The latter of keeping the other values on their 'mean' level always holds but is not mentioned in the upcoming explanation of the effects of not being repetitious.

Information can also be obtained by comparing probability values between covariate categories. An observation is that males have a higher probability value for the (1) C class than females. As the ProbMeans output shows probabilities, the following reasoning can be thus made: conditional on being a male or female, men are compared to females more likely to be present in the (1) C class compared to the other classes. However, the difference in probabilities does not give information

Table 17. ProbMeans 5-class solution: outside-class probability distribution from the latent class model with indicators and covariates; covariates are only shown.

Class[a]	(1) C	(2) B	(3) CW	(4) CB MM	(5) PT+ MM
Class size (%) N = 183,618	41	20	17	12	10
Covariates					
Socio-demographic variables					
Gender (%) (Wald = 792, p < 0.00)					
Male	47	18	15	11	9
Female	36	21	19	13	10
Age (%) (Wald = 705, p < 0.00)					
18-32	38	19	12	10	22
33-44	47	16	15	14	8
45-54	46	18	15	14	7
55-64	41	22	17	13	7
65+	35	24	26	11	5
Ethnicity (%) (Wald = 815, p < 0.00)					
Native	42	20	17	13	8
Immigrant	39	18	19	7	18
Missing	30	26	12	2	30
Education level (%) (Wald = 919, p < 0.00)					
No education	35	22	27	4	12
Primary education	34	25	28	7	7
Pre-vocational education	40	23	20	11	7
Vocational & higher secondary education	44	18	16	12	10
Higher vocational & university education	41	18	15	14	11
Missing	39	22	20	11	9
Household composition (%) (Wald = 180, p < 0.00)					
One-person household	32	24	22	8	14
Couple household	42	21	19	12	6
Other household	45	17	14	14	11
Occupation (%) (Wald = 3892, p < 0.00)					
Employed	48	17	13	13	9
Student	17	26	6	8	43
Other occupation	34	24	25	11	6
Missing	39	23	16	13	9
Household income (standardised) (%) (Wald = 341, p < 0.00)					
< €20,000	35	24	21	9	12
€20,000 - €30,000	43	19	16	13	9
€30,000 - €40,000	45	18	14	14	9
≥ €40,000	47	15	14	14	9
Missing	26	28	17	7	21

Table 17 (continued). ProbMeans 5-class solution: outside-class probability distribution from the latent class model with indicators and covariates; covariates are only shown.

Class[a]	(1) C	(2) B	(3) CW	(4) CB MM	(5) PT+ MM
Mobility resource variables					
Licensure (%) (Wald = 2955, $p < 0.00$)					
No	10	34	24	5	26
Yes	46	18	16	13	8
Missing	15	32	44	1	7
Ownership household car (%) (Wald = 5225, $p < 0.00$)					
No	5	40	23	3	29
Yes	46	17	16	13	7
Ownership company car (%) (Wald = 1600, $p < 0.00$)					
No	40	21	17	12	10
Yes	70	4	14	11	2
Ownership household bicycle (%) (Wald = 2569, $p < 0.00$)					
No	44	5	35	3	12
Yes	41	21	16	13	10
Built-environment variables					
Urban density (%) (Wald = 1953, $p < 0.00$)					
High (≥ 1500 addresses/km^2)	35	22	18	10	15
Medium (1000-1500 addresses/km^2)	43	20	17	14	7
Low (< 1000 addresses/km^2)	48	17	16	14	5
Time					
Year (%) (Wald = 290, $p < 0.00$)					
2010	41	18	19	13	9
2011	42	19	16	13	10
2012	42	20	16	12	9
2013	41	19	17	12	10
2014	40	21	16	13	9
2015	41	20	17	12	10
2016	41	20	18	11	10
2017	42	20	17	11	10
Inactive covariates	Appendix C; latent class membership.				
Multimodality measurements of travel behaviour					
Built-environment variables					
Residential municipality (%)					

Note: Some row values may not add up to 100% due to rounding.
a: (1) C: Car exclusive, (2) B: Bicycle mostly, (3) CW: Car + Walk, (4) CB MM: Car + Bicycle Multimodal, (5) PT+ MM: Public transport+ Multimodal.

about the strength or size of the difference in probabilities of people belonging to a class. The profile output is more informative on this but has another viewpoint (Appendix C; latent class profiles). So, the observation between covariate categories among classes in the ProbMeans output does not mean that conditional on being in the (1) C class, it is likely that males are more present in a particular class than females. This type of reasoning can only be used when looking at the Profiles output (Appendix C; latent class profiles), where, by chance, also a higher share of males can be seen in the (1) C class, conditional on being in the (1) C class. Moreover, in the profile outcome, a different share of particular categories can also (partly) be explained by the different prevalence of individuals in specific covariate categories.

Based on Table 17, the association of covariates are explained, grouped per overarching variable group.

Socio-demographic variables

The associations between the socio-demographic variables and the latent classes are explained.

Males (covariate gender) are more associated with being in the (1) C class (47% prob.) than the other classes, where the probabilities decline for the upcoming classes, comparable due to lower class sizes. However, females are also more associated with being in the first class than the other classes (36% prob.). When comparing (fe)males, males are more likely to be present in the Car exclusive class compared to females, and females are more likely to be in the (2) B and (3) CW classes, as well as in the (4) CB MM and (5) PT+ MM class than males.

Concerning age, all age categories show higher probabilities in the classes with higher class sizes, or almost. An exception is for the age category 18-32, with a relatively high probability value (22%) for the (5) PT+ MM class. This value is also relatively high compared to the probability values of other age categories in this class. This finding is also reflected that the occupation student has a high probability value in this class (43%), as many in this age category are students. No other substantial differences in the distribution of latent class probability values occur when comparing the age categories. Nevertheless, it can be seen that, for instance, the age category 65+ has a relatively high probability, compared to the other age categories, to be in the (2) B or the (3) CW class and a somewhat lower probability to be in the (1) C class. Moreover, for the more middle-aged categories (33-64), higher probability values are present for the (1) C class, compared to the younger or older age categories.

The ethnicity of people, being from native or immigrant origin, shows overall similar probability value proportions compared to the class size. One finding is that someone with an immigrant origin has a slightly higher association with the (5) PT+ MM class (18% prob.) and a slightly lower association with being in the (4) CB MM class (7% prob.), also compared to the native origin. It could be that immigrants tend to live more often in dense urban areas, as those are mainly more culturally diverse. Moreover, dense urban areas have better public transport connections (see 5.4).

When looking at the education level, most associations per category with the latent class probability distributions are the same as one could expect based on the class sizes. Some differences between education categories are visible. The category vocational education up to university education has higher probability values than the other categories for the (1) C and the (5) PT+ MM class. On the other hand, they have lower probability values for membership in the (2) B or (3) CW class. Individuals with no education or primary education are less associated with the (4) CB MM or (5) PT+ MM class but somewhat more with the (3) CW class compared to the other categories.

The household composition shows that the latent class probability distribution is comparable to the expectation of following the class size. At the same time, some differences between the categories are visible. One-person households have a lower probability value for the (1) C class than couple households or other households, meaning that those are less likely to be in the (1) C class. However, they are more likely to be in the (2) B, (3) CW and the (5) PT+ MM class. The other household category, having other compositions and mainly three persons or more in the household, is less likely to be in the (2) B or the (3) CW class compared to the other categories and has the highest probability value for the (1) C class.

Regarding occupation, as mentioned, students are highly likely to be in the (5) PT+ MM class compared to the other classes and compared to the other categories, with a probability value of 43%. The other distributions over classes of probability values are comparable to the class size. Moreover, students are also more likely to be in the (2) B class and less likely to be in the (3) CW class and the other multimodal class (4) CB MM compared to the other classes and the other categories. Individuals with 'other' occupations (individuals who are working in their own household, are unemployed, are unable to do work or are retired) are more associated with the (3) CW class than the other categories. At the same time, employed individuals are less likely to be in the (2) B class compared to the other categories and most likely to be in the (1) C class compared to the other classes.

The household income shows, like most variables, a latent class probability distribution which follows more or less the class sizes. But, the higher income categories are somewhat more associated with being in the (1) C class, compared to the lowest income category (less than €20,000). On the other hand, the lowest income category is more associated with the (2) B class and the (5) PT+ MM class than the other categories.

Mobility resource variables

The mobility resource variables and their effects on class membership are elaborated.

The licensure and ownership of a household car variables show that for individuals with a car licensure and/or individuals who own a household car (category yes), there is a higher latent class probability value for classes with a larger class size (46% prob.). This seems plausible, as the highest class size (1) C has the highest car use, and the lowest class size (5) PT+ MM has the lowest, where the classes in-between vary somewhat in car use. This distribution is clearly not the case for the no category of both variables. Individuals not owning a car license and/or not owning a household car is more likely to be in the classes with less car use, namely (2) B, (3) CW and the (5) PT+ MM class (≥ 23% prob.). Although the (3) CW class has some car usage (as a passenger), the probability of belonging to the (4) CB MM class with quite some car usage is relatively low (about 5% prob.). The difference between the yes and no category in both variables is also quite substantial in all the classes, meaning that, e.g., having licensure has a greater probability of belonging to the (1) C class than having no one.

The ownership of a company car also shows a high probability of being in the (1) C class, with a probability value of 70% that the individuals belong to the (1) C class, which is very high compared to the other probability values among the classes. The difference in this (1) C class with the no category value is also relatively high. However, many who do not own a company car also have a high probability (40%) of belonging to the (1) C class compared to the other classes, as, in general, ownership of a household car is high among the sample. The prevalence of individuals owning a company car in the (3) CW (14% prob.) is also higher compared to the resulting classes with hardly any car use, the (2) B class and the (5) PT+ MM. When not owning a company car, the probability distribution among the classes is more comparable to the class size, and for classes (3) CW and (4) CB MM, almost the same probability values occur compared to the yes category.

When owning a household bicycle, a comparable probability distribution is seen compared to the class size. Still, a high likelihood exists for individuals to belong to the (1) C class (41% prob.), the biggest identified class, as almost all car owners also own a bicycle. However, these individuals have a higher probability than when not owning a bicycle to be in the (2) B and (4) CB MM classes, which are classes with bicycle usage. Not owning a bicycle is more associated with the (1) C and (3) CW classes compared to the other classes. Still, there is some probability (12%) for individuals with no ownership to be in the various mode usage (5) PT+ MM class, meaning that most likely walking, PT and the car are used for these individuals. Moreover, not owning a household bicycle shows a high probability value of being in the (3) CW class compared to the other classes, which means that mostly likely the car is primarily used in this class or another bicycle which is not in their ownership.

Built-environment variables

Urban density shows an overall declining probability value among the classes comparable to the class sizes. However, individuals with a high urban density are more associated with the (5) PT+ MM class (15% prob.) than those in the other classes or categories. Likewise, individuals with a low urban density show a relatively high probability value (48%) in the (1) C class compared to the high urban density category and compared to the other classes.

5.4 Spatial distribution of latent class membership

This section shows geographical visualisations of the inactive covariate residential municipality of the model outcome involving probabilities across classes (section 5.3). Inactive means that this variable is only used to describe the identified classes. First, a visualisation of urban density is shown, to which the other residential municipality explanations can be related. Second, the dominant class per municipality is visualised based on the highest class membership probability. Third, the probability of being a class member as a travel user per municipality is shown geographically for all the classes.

Urban density

First, the urban density distribution of municipalities in 2017 is shown in Figure 15. The model also includes this built-environment variable as an active covariate (section 5.3). However, the urban density is now separately visualised to see how it is distributed across the municipalities to use for the following sections. It should be noted that the municipal organisation with corresponding urban density in previous years was slightly different. For 2010-2017, only the 2017 municipal organisations and urban density were used in this visualisation. However, the urban density from individuals in the model is used from the corresponding years, which might be different than in a later year after the reorganisation (see section 4.5, inactive covariate). A crucial side note is that the urban density and residential municipality variables only show where individuals live and not where individuals are travelling. However, the characteristics of the (most likely) starting point (residential municipality) of travelling can likely influence travel patterns.

When characterising The Netherlands in terms of density, the Eastern part is more the periphery of The Netherlands, with a low urban density. This is compared to the highly urbanised Western part, also called the ‘Randstad area’, with the four biggest cities in The Netherlands in terms of population (Amsterdam, Rotterdam, The Hague, Utrecht). In general, good connections within and between these cities are the case for the public transportation network.

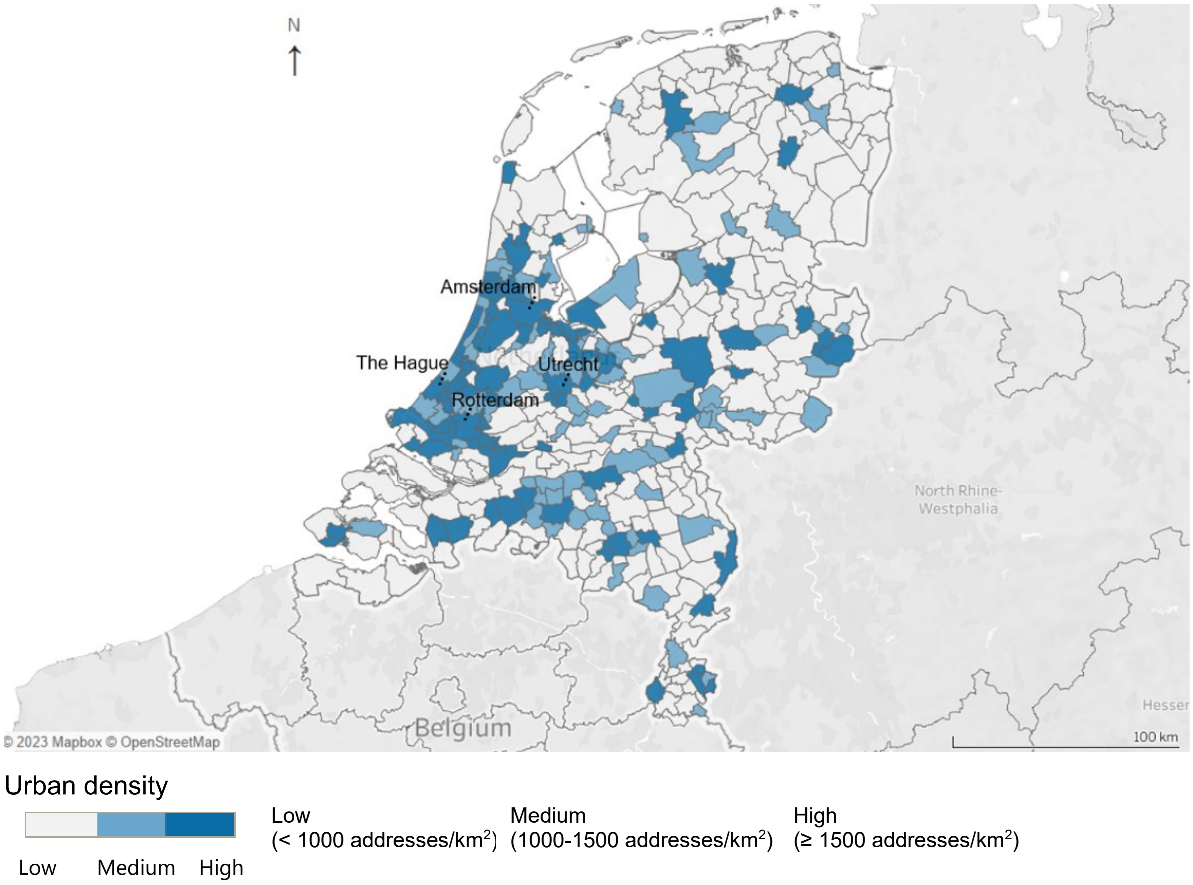


Figure 15. Spatial distribution urban density of residential municipalities in The Netherlands 2017.

Dominant class per municipality

Second, it is of interest to see if any municipalities have certain dominant classes based on the highest probability value among the classes based on the ProbMeans output (section 5.3) before analysing the probabilities per class. Most municipalities likely have the (1) C class as the dominant class, which is the biggest class. It can be clearly seen that almost all municipalities (in 2017) have the (1) C as the dominant class, based on the highest distribution probability value (Figure 16). Some deviations from this can be seen based on geographical visualisation. However, only some Wadden islands (1, 6 and 9), Groningen (4) and Amsterdam (13) have a dominant value, which is not close to a probability value for the biggest classes with generally the biggest probability values: (1) C class or the (2) B class (Table 18). For instance, the dominant class of Delft (2) is the 2 (B) class, but the probability value is really close to the (1) C class.

Most Wadden Islands (1, 6, 7 and 9) in the North Sea (Figure 16), with a low urban density, and some bigger cities (6 cities) with a high urban density have the highest probability of belonging to the (2) Bicycle mostly class. This might be due to the characteristics of the Wadden Islands (e.g., small size, bicycle-friendly), and for the bigger cities due to quickly reachable historical centres or having a higher student population. One other Wadden Island (12), with a low urban density, has the highest probability of belonging to the (3) CW class. Two other municipalities, which have medium and low urban densities (2 and 11), also have the highest probability of being in the (2) B class. These findings suggest that other characteristics than the included urban density in the model, like infrastructural characteristics (of the municipalities travelled in) and cultural (city) related aspects, could determine the use of the bicycle in travel patterns, besides the personal, household related and mobility resource variables determined in section 5.3. Another finding is that Amsterdam (13), with a high urban density, has an equal probability of belonging to the (2) B or the (5) PT+ MM class, where it is likely that the high urban density plays a role in using public transportation due to the wide availability of it.

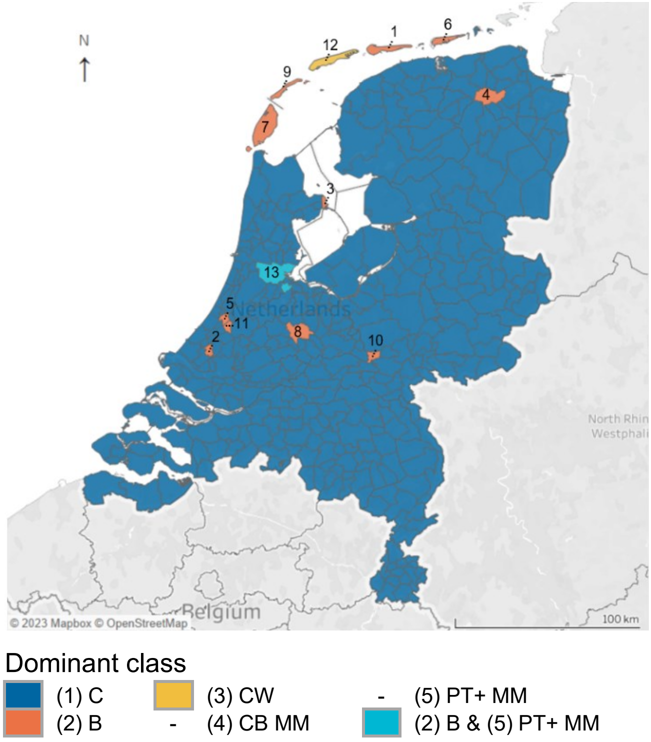


Figure 16. Spatial distribution dominant class(es) (of highest membership probability) by residential municipalities (388 2017) in The Netherlands for 2010-2017 (ProbMeans 5-class solution distribution; inactive covariate is shown).

Table 18. Dominant class(es) per residential municipalities (see Figure 16).

Municipality (NR in this figure)	**Dominant class**[a]	**Urban density**[b]	**(1) C**	**(2) B**	**(3) CW**	**(4) CB MM**	**(5) PT+ MM**
Else than below	1	-					
Ameland (1)	2	3	0.21	0.46	0.13	0.18	0.01
Delft (2)	2	1	**0.27**	0.29	0.16	0.09	0.20
Enkhuizen (3)	2	2	**0.29**	0.29	0.22	0.11	0.08
Groningen (4)	2	1	0.22	0.36	0.15	0.12	0.16
Leiden (5)	2	1	**0.25**	0.28	0.14	0.12	0.20
Schiermonnikoog (6)	2	3	0.09	0.64	0.05	0.12	0.09
Texel (7)	2	3	**0.30**	0.32	0.16	0.21	0.02
Utrecht (8)	2	1	**0.26**	0.26	0.15	0.10	0.23
Vlieland (9)	2	3	0.11	0.53	0.31	0.04	0.01
Wageningen (10)	2	1	**0.29**	0.30	0.16	0.10	0.14
Zoeterwoude (11)	2	3	**0.29**	0.31	0.15	0.17	0.08
Terschelling (12)	3	3	**0.23**	**0.26**	0.30	0.18	0.02
Amsterdam (13)	2 & 5	1	0.21	0.28	0.18	0.07	0.28

Note: Some values are the same per row due to rounding.
a: Dominant class(es) based on the highest probability value among the classes.
b: See Figure 16.
Bold = A municipality with probability values on classes other than the dominant class, which are close to the dominant class.

Probability distribution class membership residential municipality

Third, the residential municipalities can be spatially visualised per class to see the distribution of class membership of travel users from a particular municipality. As municipality is an inactive covariate, this variable only describes the founded classes consisting of indicators and covariates. These figures (Figure 17, Figure 18, Figure 19 used the ProbMeans output, where the latent class distribution can be seen per class per municipality (section 5.3). So, for instance, a value of 41% of a municipality shows that there is a probability of 41% for those residents, given their residential municipality, to belong to the (1) Car exclusive class. Suppose this is the highest value among the classes. In that case, this class has the highest class membership probability for that municipality (this is not visualised in this figure). If this value is also higher than another municipality, this municipality is more likely to belong to this class. However, these values compared between municipalities (visualised in these figures) do not exactly show the (strength) difference in effects. The advantage of using the ProbMeans output is that the probability among classes is used instead of using the dominant class per municipality (the class with the highest ProbMeans value, Figure 16), where this more careful interpretation with (un)certainty is lost.

When looking at the conditional probabilities for municipalities, one should be aware that the scales of percentages differ per map. Moreover, these findings should be interpreted with care as colour schemes are used, which can be less interpretable by the eye and differ per human. Moreover, in this case, all municipalities are represented, which reflects many colours and values to be compared and interpreted. Although, at a glance, some spatial clusters can be easily seen instead of a long list with values.

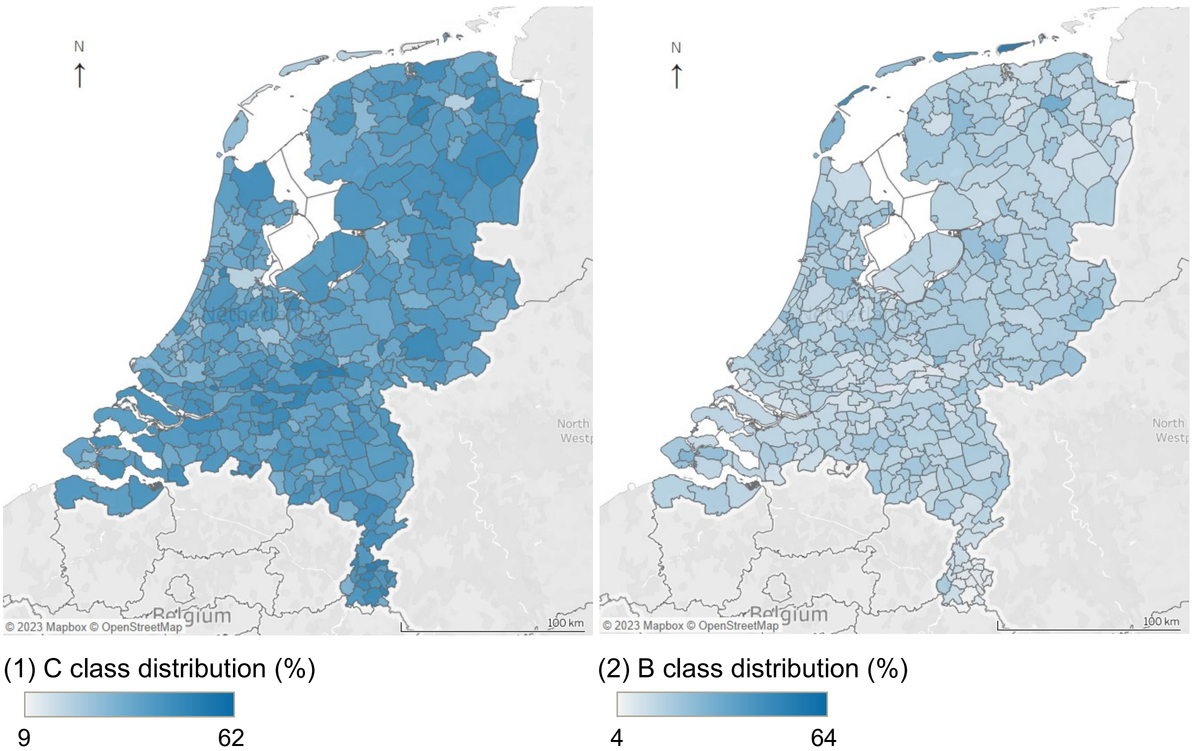


Figure 17. Spatial distribution (1) Car exclusive class (a, left) and (2) Bicycle mostly class (b, left) of membership probability by residential municipalities (2017) in The Netherlands for 2010-2017 (ProbMeans 5-class solution distribution; inactive covariate is shown).

The maps are subsequently described.

When looking at the first map, Figure 17 (a), which shows the probabilities of travel users from municipalities belonging to the (1) C class (based on the latent class probability distribution among the classes per residential municipality), there is a slight overlap of comparable values in some Eastern municipalities, in general, less urban dense areas. They generally have a higher probability that an individual with that residential belongs to the (1) Car exclusive class. Some variation across the country can be seen, and the range of percentages is 53%.

The colour map for the (2) Bicycle mostly class shows hardly any variation across the country (Figure 17, b). Unsurprisingly, residents from the Wadden Sea Islands are more likely to belong in this class. The range is 60%, which is quite extensive and can cause more difficulty seeing differences as the difference should be higher to see them correctly.

The (3) Car + Walk distribution shows that some southern travel users of some municipalities have the highest probability of belonging to this class (Figure 18, a). However, it is difficult to see other variations, while the range of values is relatively small (31%).

The (4) Car + Bicycle Multimodal distribution shows more variety across the country, although this can also be due to the small range of percentages shown (18%). Still, a clear pattern is not visible (Figure 18, b). However, travel users from Rotterdam and Amsterdam (large and dense urban cities, Figure 15) have a lower probability of belonging to this class compared to the other residential municipalities.

As already discussed, Amsterdam had a higher and the same probability of belonging to the (2) B and (5) Public transport+ Multimodal classes. Moreover, the most multimodal class (5) clearly shows that the Western highly urbanised 'Randstad area' has a higher probability of belonging to this class than the other parts of The Netherlands Figure 19. The range of these values is 28%. As the classes are characterised in terms of patterns and determinants (sections 5.2 and 5.3), the next part is to see if the patterns are developing (section 5.5).

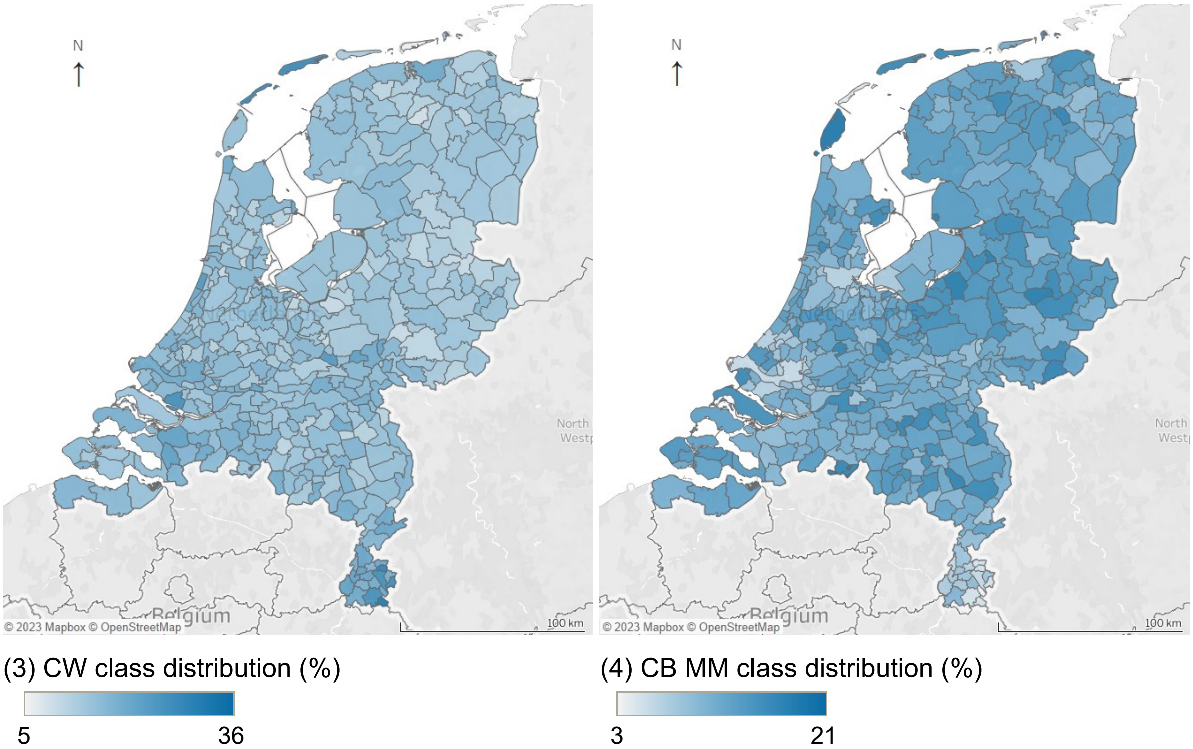


Figure 18. Spatial distribution (3) Car + Walk class (a, left) and (4) Car + Bicycle Multimodal class (b, left) of membership probability by residential municipalities (2017) in The Netherlands for 2010-2017 (ProbMeans 5-class solution distribution; inactive covariate is shown).

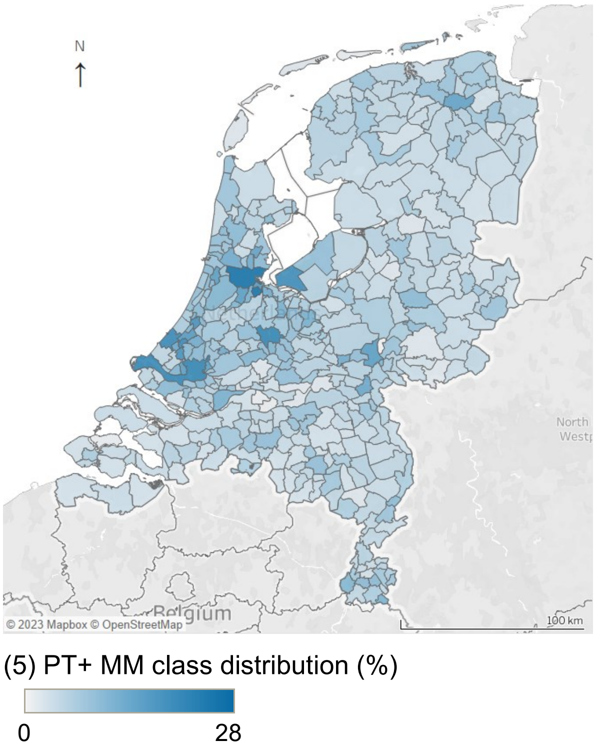


Figure 19. Spatial distribution (5) Public transport+ Multimodal class of membership probability by residential municipalities (2017) in The Netherlands for 2010-2017 (ProbMeans 5-class solution distribution; inactive covariate is shown).

5.5 Development of latent class membership and multimodal travel patterns

Figure 20 shows the earlier identified travel patterns captured in travel users' classes (sections 5.2 and 5.3) over time in several colours. Per travel pattern (class), the latent class distribution is shown conditionally on the covariate year (ProbMeans output) to show the development of the identified travel patterns of travel user classes.

So, for instance, conditional on the year 2010, a probability of 41% means that an individual travel user belongs to the (1) C class (dark blue line) as a Dutch resident, and conditional on the year 2011, there is a probability of 42% that the individuals from belong to this class. This shows that, as an example, in 2011, there was a slightly higher probability of belonging to the (1) C class compared to 2010. The (1) C class has an average class size of 41% over the years (Table 16), shown in the dotted lines. This is done for every class to easier interpret the development over time (via the conditional distributions).

The development of the class sizes shows several findings (Figure 20). The (1) C, (4) CB MM, and (5) PT+ MM class are all sometimes stable and sometimes show slight decreases or slight increases compared to the previous year and/or compared to the average class size. However, the range of probability values across all years is only 2% or smaller for those classes. This means that the class with the lowest degree of multimodality (1) C and the classes with the highest degrees of multimodality (4 and 5) are quite stable in terms of class size. Interestingly, the (2) B class shows a range of 4%, which slightly increases from the lowest point (in this timeframe) in 2010 towards the highest point in 2014, after which a slightly decreasing trend is set towards the mean class size. For the (3) CW class, a decrease equal to 3% is visible from the highest to the lowest point from this period, from 2010 to 2011, after which the pattern continues in a stable way, slightly deviating from the mean.

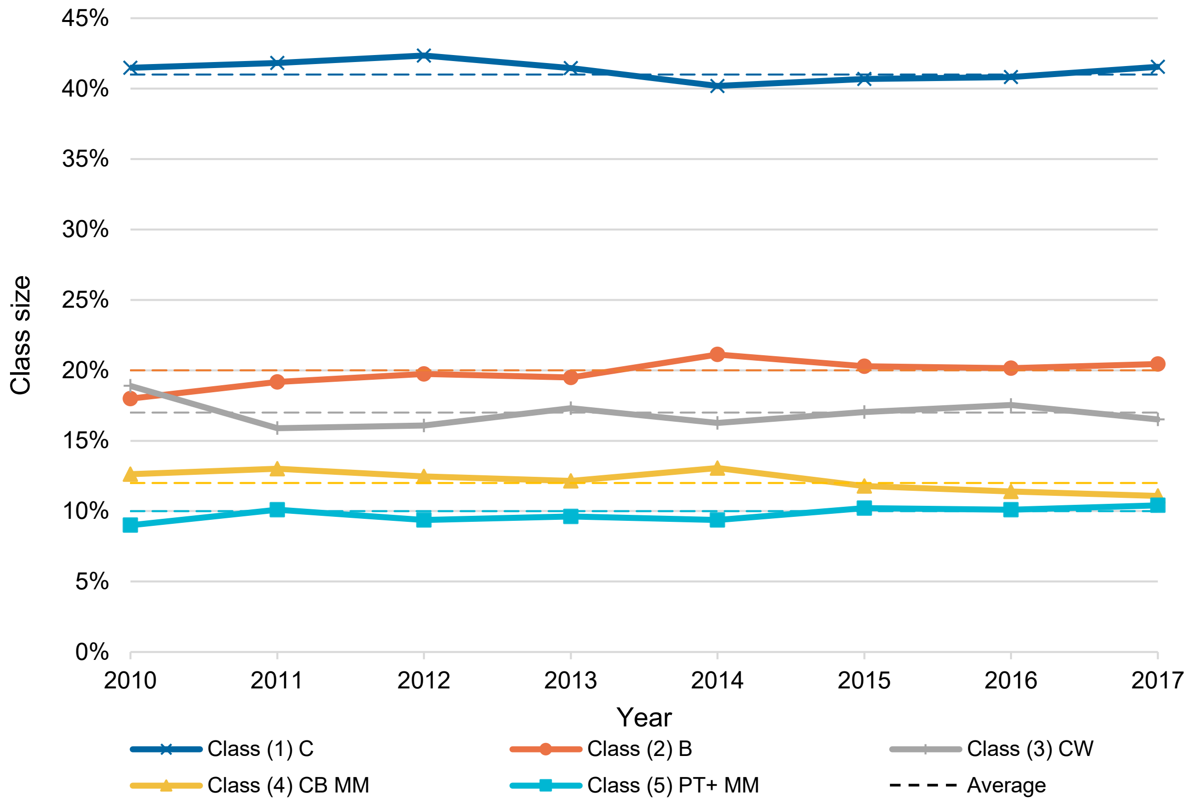


Figure 20. Class membership conditional on year (ProbMeans 5-class solution distribution; covariate is shown).

The depicted visualisation over time (Figure 20) explicitly showed the class sizes' development. However, implicitly, it reflects the development of the classes, which all have multimodality measurements and are characterised by modal travel patterns. But, no (linear) trend can be identified regarding the overall group sizes; thus, the same holds for the multimodal travel patterns of the travel user groups. This is comparable to earlier findings of the stable overall trends of travel-related determinants, variables and multimodal measurements (section 5.1). So, overall, class membership conditional on year showed that the probability of belonging to a latent class is, to a great extent, stable from 2010 to 2017 in The Netherlands, where the possible shifts of individuals (with specific characteristics) between classes are not visible in our results. In other words, the Dutch mobility system shows overall no developing modal travel patterns in terms of more or less practised specific behaviour overall.

It should be kept in mind that the distribution within each class and the distribution of the multimodality measurements per year is not identified in this model, which might have had differences between years due to individuals shifting from one class to another. Some notions are explained.

So, this method does not allow for looking at the change of determinant's influence per year on the probability of individuals belonging to a particular class reflecting specific modal travel patterns. Nevertheless, a prudent assumption can be made that based on the shown trend of class sizes, and the stability of the input variables, it is likely that no major changes in the effects of determinants, in this aggregate view, occurred in this period.

Moreover, although the input variables for the multimodality measurements are highly stable, it is not expected that these fluctuations influenced the measuring. As Heinen & Mattioli (2019a) explained, if an individual's number of trips or stages decreases, this affects the multimodality measurement. However, the correlations between the multimodality measurements and the number of stages were only moderate in their study, indicating that their decline in the number of trips had a negligible effect on the reducing level of multimodality.

To place these findings into context, it is remarkable that the emerged (multi)modal travel patterns were highly stable for the period of 2010-2017 (in The Netherlands). Given the changes in demographics, which might exhibit a transition towards other modes, and unavoidable sample fluctuations in terms of individuals and representativeness (year is not reflected in overall representativeness in section 4.8), the modal travel patterns are hardly developing and highly stable over the examined eight years. This prominent notion shows that the increasingly known urgency of climate change, (positive) changing attitudes and awareness of individuals and practitioners of the last decades (section 1.1) towards the urgency and impacts of climate change have not accomplished the desired effects of creating more active, sustainable (and multimodal) travellers.

6. Conclusion and Discussion

The effects of global warming are becoming more severe, like weather or climate extremes, and are impacting our lives. One of the main contributors to global warming are the emitted greenhouse gas emissions by human activities. Especially the transport sector has a high cause of this. Moreover, in this sector, emission growth has continued over the last decades. Regarding passenger traffic, road traffic (by car) is the main cause of high energy consumption and emission. Due to car dependency in industrialised countries, the travel behaviour of passengers remained unsustainable. The high car dependency in countries also played an essential role in limited shifts to more carbon-friendly alternative modes.

Researchers and practitioners proposed that increasing the multimodal travel behaviour of travel users is a potential way to stimulate more sustainable mode usage, implicating social relevancy such as limiting environmental impacts and improving public health and overall prosperity. Generally, multimodality is defined as a diverse (and balanced) mode usage, which means that someone has the highest level of multimodality when a variety of modalities are used, with about equal intensities, in a specific time period. Multimodality is also referred to as (modal) intrapersonal variability, a research stream examining the variety of modes used within an individual's travel patterns.

In previous studies, it has been found that multimodal travellers are generally more sustainable (emitting less CO_2) than monomodal car users. Especially the notion of comparing to strict car users (travel users only using the car) is important, as multimodal travel behaviour can also involve car usage. Moreover, using only one or two modes from the mode set bicycle and walk, and thus having monomodal or bimodal behaviour, is more sustainable than multimodal behaviour. But interestingly, multimodal travellers are more likely to change their behaviour to more active or sustainable modes when the right conditions are provided, making multimodal travel behaviour a fascinating phenomenon. Within multimodal behaviour, the use of active travel modes (requiring a physical effort) like walking and cycling has a unique role. Besides the carbon-friendliness of these modes, it is argued that a strict car user is less likely to switch to public transport, in contrast to a car user who already occasionally uses a bicycle, for instance, as it can be easily used as an access and egress mode for public transport.

Among practitioners, improving multimodal behaviour is also regarded as a potential first step in moving away from habitual car use, for instance, by promoting the use of more sustainable and active travel. Governments are working more often on integrating active and sustainable mobility in the transport system by adopting mobility plans and resolutions, showing commitment and improved awareness. However, the high car dependency is still visible, in general, in a risen private car ownership and company car (made available to employees by the company) ownership. It is observed that multimodal travel behaviour is only limited visible in countries.

Several aspects remain to be explored in the interest of researching multimodal travel behaviour. First, the literature has not yet agreed upon how to measure multimodality and use the different aspects of measures for specific research purposes. Second, socio-demographic (including household-related), travel mode availability and urban context determinants of multimodal behaviour showed some critical factors in previous research. Despite the importance of including mobility resource variables in research and the known car dependency, hardly any study involved company car ownership as a determinant in multimodal behaviour research. Third, studies investigating trends in measured multimodality are scarce, but, more importantly, the development of corresponding modal travel patterns, including various combinations of mode usage, remains to be determined.

Latent Class Cluster Analysis (LCCA), also named Latent Class Analysis (LCA), allows for measuring multimodality whilst enabling capturing modal travel patterns themselves. LCA identifies travel groups consisting of travel users practising specific patterns. When referring to multimodal patterns, it is about analysing multiple mode usages in travel patterns, and the degree of multimodality is about how multimodal those patterns are. More specifically, LCA can capture comparable travel patterns of individual travel users and clusters them into homogenous groups. These groups, or so-called classes, are unobservable in real life, as the assumption is that certain latent classes exist in society, which can be made emergent with the use of an LCA. Methodology-wise, a time variable in LCA is included to analyse the developing size of the classes reflecting multimodal travel patterns. More targeted policies can be created by understanding the types of travel users over time and aspects that play a role in having certain multimodal travel patterns. Specific comparable classes with comparable modal patterns with a certain degree of multimodality can thus be targeted differently.

The identified knowledge gaps and the suitable method of analysing comparable groups resulted in the following main research question (MRQ): "*How are travel user groups comprising multimodal travel patterns characterised, and how do determinants and time influence the travel behaviour?*" to analyse the determinants and development of characterised multimodal travel patterns.

The multimodal travel patterns were once identified for individual adult travel users doing daily travel in The Netherlands from 2010 to 2017, and insight is given in the class sizes per year with LCA. Cross-sectional national data from the Dutch National Travel Survey (OViN), which Statistics Netherlands administers, is used. This is a high-quality data set in which many individual participants filled in a one-day travel diary about the trips they made, including their personal and household characteristics. The modalities walking, cycling, public transport (bus/tram/metro/train), and car (driver or passenger) are used to define the travel patterns. Individuals are only included when doing daily travel in The Netherlands for the whole day by only using modes from the defined modes set (their full travel pattern). The previous is because including parts of travel users' travel patterns results in biased and incomplete travel patterns. The level of analysis is at the stage level, as it accounts for the variation in all possible modes used in a time period. Stages are part of a trip, e.g., a trip can comprise a walking, public transport, and cycling stage. So, the number of stages per travel mode is used as a measure of intensity for mode usage in the LCA, which groups individuals based on this measure.

This chapter consists of the conclusion (section 6.1) to answer the research questions and provide the research contribution, the policy and societal implications (section 6.2) and the discussion with future research (section 6.3).

6.1 Research questions conclusion and research contributions

Subsequently, the research questions (RQ1-7) are answered, after which the main research question is answered (MRQ), and the research contributions are given.

RQ1: Which insights on measuring multimodal travel behaviour can be derived from previous studies?

The first step in analysing multimodal travel behaviour was to perform a literature study to overview the existing multimodality measurements. By classifying the measures into measurement categories and evaluating and comparing the categories and corresponding measures, suitable measures are chosen to assess the degree of multimodality of modal travel patterns in our research.

The predefined characterisations of multimodality, part of nominal characterisations (measurement category), were regarded unsuitable for analysing travel patterns. Because a-priori, criteria are set up in which class individuals belong, e.g., individuals only using the car to a monomodal car user class, and individuals using the car and at least one other mode to a multimodal car user class. The main disadvantage is that there is no insight into differences in travel patterns between groups, although the groups can be intuitively interpreted. This a-priori approach is impossible to use in combination with our main method used (Latent Class Analysis, see RQ4), which can be seen as a data-driven classification of multimodality (part of measurement category nominal characterisations). These data-driven classifications use clustering schemes, and depending on the data, individuals are placed in comparable groups post-hoc. Because LCA is a clustering scheme, this category is inherently used and cannot be combined with using predefined nominal groups. However, the shortcoming of grouping individuals in nominal groups with LCA is that some information is lost within the aggregated groups. Still, the travel patterns on which the individuals are clustered can be shown in LCA. The last measurement category, identified in our case, consists of numeric indicators. Formally, these indicators can be added to the LCA method to capture the multimodality of individuals within the groups. The advantage of using them jointly is better capture the individual degree of multimodality.

A downside of joint use is that the indicators are less intuitively interpreted. However, within the numeric indicator measurements category, more simple one-sided measures and more complex multi-sided measures exist in defining multimodality. By using complex measures, the modal variety and the modal intensities are both assessed to measure multimodality ('multiple sides'), which can be used next to the simpler ones (using one side) to provide a more comprehensive view of the balanced use of modes. For the more simple measures, the number of modes is chosen as it is closely related to the usage of modes among the full travel pattern and the variability of it, compared to the share of the primary mode, for instance. A higher number of modes thus indicates a higher modal variability and, thus, multimodality. Although, this measure does not take into account the modal frequencies. A measure especially suited for measuring diversity (of modes) and equality (balanced modal usage) to capture multimodality is OM_PI. The measure is based on the Shannon entropy and is called the Objective Mobility Personal Index, which is used more widely in research. So, when someone uses only one mode, this person is regarded as not multimodal and has a degree of zero. On the other hand, if a person uses all modes with equal intensity, this is regarded as the most multimodal (a degree of one).

RQ2: Which insights on multimodal travel behaviour determinants can be derived from previous studies?

An overview of multimodal travel determinants is based on previous research to define the conceptual framework for studying multimodal travel behaviour. A systematic review analysed articles using none or a variety of multimodality measurement categories (def. in RQ1). Although all studies used different measures, many comparable findings exist on which variables are likely to determine multimodal travel behaviour. Most research involved grouping or clustering individuals and examining the comparable groups of individuals by assessing the characteristics of several travel user groups. However, some research results of determinants remain contradictive, are differently operationalised, do not give insight into the travel patterns themselves, have no crystallised effects, or are examined in other contexts. Other determinants are yet not often researched, and others are necessary for analysing multimodal travel patterns.

Several determinants are thus set up in analysing multimodal travel behaviour in the upcoming RQs. The incorporated socio-demographic variables are gender, age, ethnicity, education, household composition, occupation and household income. The included mobility resource variables are licensure, household car ownership, company car ownership, and household bicycle ownership. The built environment is analysed for the variable urban density.

RQ3: What are the observed trends of travel behaviour determinants and travel behaviour?

The trend overview of potential determinants and travel behaviour-related variables were shown to put the upcoming results into context.

The used variables for determinants (RQ2) of the Dutch sample were little subject to change over the years in general. Moreover, it is assumed that the sample is highly representative to the Dutch population for the pooled years (2010-2017) by comparing our sample distribution values with the one for the target population. Besides, it is assumed that the value and distributions within the years are also representative, as limited major shifts in demographics or other aspects likely occurred.

The travel behaviour-related variables (measured over one day) are analysed for the average number of modes, the average total stages, and the average stages per mode. These variables were consistent over the years. Besides, the multimodality measurement OM_PI (Objective Mobility Personal Index), which is dependent on the number of modes used and the intensity of use (stages per mode) to analyse the diversity and balanced mode use (high multimodality), showed a stable trend.

These observed trends show that the overall (multimodal) travel patterns are hardly subject to change over the years. Although, if overall averages or distributions are similar, it should be noted that shifts could have occurred within the distributions and the same aggregate values remained. Based on this observation, it is interesting to see, unless average comparable descriptives, which 'hidden' (multi)modal travel user groups exist and if the classes have developed over time, which is investigated in the upcoming explained research questions.

RQ4: What are the captured multimodal travel patterns and degrees of multimodality by the identified travel user classes?

The main method used, Latent Class Analysis (LCA), identified five multimodal modal travel classes. This means that society involving travel users can be captured with five existing (hidden/latent) classes in society, in our case, to which comparable individuals belong in terms of their overall multimodal travel pattern. The number of classes to project society on is chosen based on several (statistical) criteria. Mainly, the improvement in model fit compared to the baseline model (society as 1 class) stagnated with a model of more than five classes. It is also shown that the classes are indeed emergent, as no class is comparable to society overall.

The class profiles are typified in the following. A Car and Bicycle Multimodal (CB MM) class is found, which involves a dominated car use, complemented by bicycle and some walking. Also, a Public transport plus Multimodal (PT+ MM) class, which involves public transport, walking, bicycling and limited car use, is found. Besides two multimodal classes, a strict car user class, named Car exclusive (C), is found. Other classes were a Car and Walk (CW) class and a Bicycle mostly (B) class. The CW class involved only car use and walking, and the B class mainly involved bicycle and limited walking use. The founded classes are quite similar in mode diversity compared to other founded classes of research in the Dutch scope (e.g., Ton et al., 2020). In terms of the degree of multimodality, the PT+ MM class, which accounts for the lowest amount of individuals (class size of 10%), has the most diverse mode usage and the highest multimodality measurement (OM_PI), closely followed by the CB MM class (class size of only 12%) for both measures. On the other hand, the C class, which reflects many individuals (class size of 41%), has the lowest multimodality measurements as this class reflects unimodal behaviour. The CW (class size of 12%) and B (class size of 17%) classes showed multimodality measurements in the middle compared to the others, where the CW class is regarded as somewhat more multimodal than the B class. Last, it is shown that the ranking of the multimodal measurements was the same for both indicators, meaning that the more intuitive (number of modes) and more complex (OM_PI) measures show equal findings about degrees of multimodality.

RQ5: To what extent do determinants influence class membership of travel users?

In the identified multimodal classes with their corresponding names (RQ5), the relationships of determinants on the probability of being a member as a travel user of a particular class are assessed. All included variables (RQ2) are assessed and showed significant results, meaning they significantly affect latent class membership. The strength of the probability of belonging to a specific class, conditional on the determinant value, is the highest for the effects of ownership of a household car, ownership of a household bicycle, and ownership of a company car. This means that the effect of the probability of being a class member with a specific variable as an individual is relatively strong compared to the other determinants. However, it should be kept in mind that the mobility resource variables are not fully regarded as exogenous (only the ownership determines the travel use and not the attitude towards a mode which determines the acquiring of ownership), so the effects cannot be regarded as fully accurate, this is reflected upon in section 6.3. Other (somewhat) strong associations are found between latent class membership and occupation and urban density. Besides our explained and founded relations, it is mainly found that the highest probabilities values of class membership are for the classes with the highest class size, as most belong to that class.

Generally, the following is found for the socio-demographic, built-environment and mobility resource variables.

It is found that males, and individuals with higher incomes, are more likely to be in a car-dependent (C class), while females are likelier to be in a bicycle mostly (B) class, and females and individuals with lower incomes are likelier to be in the multimodal (CB MM, PT+ MM) groups compared to each other. Moreover, younger age groups are likelier to be in the multimodal groups (CB MM, PT+ MM) compared to other age categories. Concerning middle-aged groups, they are more likely to be in the C class than the others. Contradictive to earlier research, our study showed that the oldest age group (compared to the others) and individuals with other occupations (mainly retired and unemployed) are not car-dependant but also not likelier to be multimodal than the other individuals, but more associated with the CW class. Regarding the student occupation, many are likely to be in the CB MM or the B mostly class compared to the other classes. This study also shows that people with an immigrant ethnicity (compared to natives) are likely to be in the PT+ MM class, possibly due to many living in dense urban areas compared, as individuals living in urban areas compared to rural areas are likely to be more multimodal.

As the first set of socio-demographic variables (and the build-environment variable) is explained, the second set of socio-demographic variables is explained. Similar findings exist based on previous research about being higher educated and being more associated with car-dependant classes (C class in our case) but simultaneously being associated with the multimodal classes (CB MM, PT+ MM). Individuals with a higher education level (compared to lower levels) are thus likely to be in one of the two (heavily) different classes. Moreover, households with a lower number of members (one-person households) have a lower probability value for the C class than couple households or other households, meaning that those are less likely to be in this class. On the other hand, the other household category (mainly involving three persons or more, has the highest probability value for the (1) C class compared to the other classes.

Regarding the mobility resource variables, having a (car) licensure, household car or no household bicycle is associated with being in car-dependant classes, consistent with earlier findings. A new finding is that company car owners are prevalent in the car-dependant class. On the other hand, when having no car or licensure, individuals are highly likely to be in the PT+ MM class.

RQ6: What is the spatial distribution across municipalities of class membership of travel users?

The residential municipality is added to the latent class model to describe the founded classes (RQ4/5). So, the probability of belonging to a particular class, conditional on the residential municipality of a travel user, is assessed. The spatial distribution of class membership probabilities is visualised in maps of The Netherlands five times (one for each class). Some spatial concentrations of higher or lower probabilities of belonging to a particular class are found in some maps. However, identifying these by eye is mainly an arbitrary task, but it can provide some general insights into which parts of the country have higher probabilities of class membership.

Generally, we identified that for the most multimodal class, the PT+ MM class, a clear pattern is visible of having higher probabilities for individuals residing (not necessarily travelling in/from) in Western urban dense municipalities. In contrast, a slightly visible pattern is observed that individuals residing in Eastern-oriented (less urban) municipalities have higher probabilities of belonging to the C class. For the other classes, more equally spread distributions are visible.

RQ7: How do travel user class sizes and classes' multimodal travel patterns develop over time?

The model outcome (RQ4/5) shows conditional on the year (2010-2017) the class size of every identified class. The class sizes are already explained (RQ4) and reflect how many individuals are practising specific multimodal travel patterns on average for 2010-2017 in The Netherlands. So, the development is reflected by the patterns and their degree of multimodality.

It is revealed that some slight fluctuations of class sizes around the mean class sizes are seen, and some classes show a declining or increasing trend for some sequential years, which is regarded as a stable pattern as the range of values per class is at a maximum of 4%. Comparable to the more general travel behavioural trends (RQ2), which were consistent over the years, the travel pattern distribution among society has been tremendously stable.

MRQ: How are travel user groups comprising multimodal travel patterns characterised, and how do determinants and time influence the travel behaviour?

In The Netherlands, using pooled data from 2010-2017, our results show that the mobility system consists of mainly car-dependant users (41%), characterised by the car-exclusive class, and many individuals are characterised by belonging to the bicycle mostly class (20%). The midst-prevalent class is the Car and Walk class (17%). The class sizes of the most multimodal classes, the Car + Bicycle Multimodal and the Public transport+ Multimodal (PT+ MM) classes, are the lowest, with around 12% and 10%. They are characterised by the highest multimodal measurements (the number of modes and OM_PI, the latter reflecting a diverse and balanced mode use).

The probabilities of belonging as an induvial to a particular class, assessed pooled-wise for the whole period, show that several socio-demographic, mobility resource and built-environment variables are determinants of class membership. The effects were all significantly affecting latent class membership.

However, the effects of ownership of a household car, ownership of a household bicycle and the ownership of a company car are the strongest. However, they are likely, as well as of licensure, in reality, somewhat lower due to endogeneity (see RQ5). Moreover, it is found that, in this study, more multimodal behaviour is in general prevalent among females, individuals with immigrant origin, younger age groups and students, some higher educated individuals, individuals with lower incomes, households with fewer members, those not having a car licensure, not owning a household car or company car and owning a household bicycle. The spatial distribution showed that, especially for the PT+ MM class, higher probability concentrations were visible in the urbanised Western part compared to the peripheral part, comparable to the finding of individuals with a higher urban density being more associated with multimodality. In contrast, the spatial probability distribution for the other classes showed less precise results.

The class sizes of the identified travel user groups of travel patterns barely develop, which remarks that the multimodal travel patterns of society, captured by the classes, are not developing over time. Given changes in demographics, fluctuating sample representativeness, and the increasing awareness of the impact of climate change among individuals and practitioners, it has not accomplished the desired effects of creating more active, sustainable (and multimodal) travellers. To conclude, it is found that mobility patterns in 2010-2017 in The Netherlands are hardly subject to change, despite changing awareness of environmental impacts and having sustainability higher in the political agenda compared with previous years. This could indicate room for improvement of policies, which is explained in section 6.2

Research contributions

Besides our research results in this section, several contributions have been made. Our results involved the development of multimodal travel patterns of classes involving comparable travel users, in which multimodality measures are added to assess the degree of multimodality.

We have shown that complementing multimodality measures can advance their use separately regarding complexity versus an intuitive interpretation. Nevertheless, most measures showed quite similar results (in terms of ranking of the classes), which is again emphasised, compared to the existing body of knowledge. A starting point of classifying and grouping the benefits, drawbacks and characteristics of multimodality measures is made, which can lead to a better understanding of how to measure multimodality effectively.

Furthermore, our analysis added to the scientific discourse in determining the potential effects of variables belonging to certain travel user classes comprising multimodal travel patterns. As many researchers identified the profiles with characteristics of specific classes involving multimodal travel patterns, we intriguingly showed the effects of determinants on the possibility of belonging to one of the identified classes. We especially revealed that including the company car variable is very important in analysing multimodal travel patterns.

Moreover, we have characterised the multimodal classes for the whole Netherlands, which has not been done before, by visually showing the distributions of the likelihood of belonging to a particular residential municipality which gave information on where (geographically) more or less multimodal travellers reside.

Last, this research contributed by tracking multimodality temporally for several mobility groups, as few researchers had shed light on this. We acquired a more comprehensive view of measured multimodality and the corresponding multimodal travel patterns. The new empirical finding of very stable travel patterns (in The Netherlands, 2010-2017) shows the importance of further researching why people keep practising the same (un)sustainable travel behaviour.

6.2 Policy and societal implications

Based on the obtained knowledge about travel user classes (in society) and their multimodal travel patterns and degrees of multimodality (section 6.1), policies can be targeted specifically to the classes to stimulate more multimodal behaviour, by influencing travel users' mode choices to achieve limiting environmental impacts, improving public health and overall prosperity. Policy directions are identified based on existing mobility plans and provided for three different perspectives (governments, public transport/IT operators, and employers). Although multimodality is defined over a more extended time period, most policies focus on improving multimodality in one trip, which also improves overall multimodality.

First, several governmental departments and cities are working on improving sustainable mobility involving active travel modes and public transport by identifying possibilities or creating initiatives to integrate multimodal hubs and Mobility as a Service (MaaS) in the multimodal mobility system. However, the initiatives are spread among diverse public-private parties (Rijkswaterstaat Duurzame Mobiliteit, n.d.; Witte et al., 2021). So, this policy direction is referred to the 'government' as one general institute, as it is suggested that the government can play an important role (Witte et al., 2021), leaving room for which specific departments could be suited for incorporating policies.

Multimodal hubs are physical places among roads (mainly outside the city ring) where the travel user can change modes, for instance, from the car to the (shared) bicycle of a (shared) mobility system or urban public transport network from several operators, as explained by researchers from VerhoevenCS and Rijkswaterstaat (Arntzen et al., 2020). These multimodal hubs can contribute to safe, accessible and liveable cities (Rijkswaterstaat Duurzame Mobiliteit, n.d.). Moreover, according to the KiM Netherlands Institute for Transport Policy Analysis, the hubs can reduce the switching resistance, facilitate traffic flows, cluster spatial services and facilitate shared mobility and electrification (Witte et al., 2021). The (help for) creating (or extension) of physical mobility hubs at specific places (with specific dominant identified travel user classes) could enhance or make it easier to have multimodal behaviour. By enabling to switch between several modes and services, several classes could be easier incentivised to choose a multimodal trip.

Mobility as a Service is an app to plan, book and pay for several transport options for one trip to enable door-to-door travel (Ministerie van Infrastructuur en Waterstaat, 2019a). The anonymous travel data can be used for new insights to reach emission goals, limit traffic jams and public transportation capacity pressure, and improve affordability. Rijkswaterstaat (the executive agency of the Ministry of Infrastructure and Water Management) has intermediated the parties developing Maas pilots, constituting several pilot apps to target several user groups specifically (Ministerie van Infrastructuur en Waterstaat, 2019b). The MaaS also gives a better potential for the multimodal hubs, as switching modes using MaaS could play a central role at the multimodal hub for specific travel users (Witte et al., 2021). Besides the existing intermediary role, the government could put forward softer policies to create awareness and promote using this potential service.

Second, the explained notion of multimodal hubs and MaaS emphasises switching modes to public transportation. The promotion and development (by several parties) of multimodal travel planner apps are suggested, as the already explained MaaS, which has benefits for travel user classes, is not yet fully offered to travel users in The Netherlands.

A role for the state-owned principal Dutch passenger railway operator NS (in Dutch: Nederlandse Spoorwegen) could be to enhance the use of the train modality in a multimodal trip. Other public transport operators could also play an important role. However, because of the longer distances which can be travelled by train and NS being the biggest railway operator, a policy direction is set up for the NS. They are working on further enhancing their multimodal travel planner by offering more (shared) modalities in their app (NS, 2021). Other (mainly private) apps exist made by (IT) developers in The Netherlands involving sub-modalities or a wide range of modalities (Zijlstra & Bingyuan, 2023), which could also be further developed and promoted.

Besides the public transport and IT operators, the government could play a role, like in the Maas Pilots, to use an integral approach to prevent the sprawl of applications. Nevertheless, the tension between the different goals of public and private parties in developing this and the competition between private apps needs to be considered in how realistic it is that 'one' new perfect app could be created. As many apps exist, the already identified direction of promoting (by several parties) the use of them could be a starting point for enhancing multimodal travel behaviour for travel users.

Third, the Dutch government is working on a legislative proposal to let employers (big organisations) stimulate the use of sustainable travel to work by employees, like using public transport or the bicycle (NOS, 2022). Moreover, the government is working on making specific policies to make the sustainable options employers provide more financially attractive (again) for employees (NOS, 2023). As many people tend to have day-to-day work behaviour, and mobility resource rules are already visible between employers and employees, a unique position for the employers in influencing the mode choice is already allocated.

Some other organisations are already working on creating more institutional rules or programs for their employees to stimulate the use of sustainable modes and limit car use (NOS, 2023). Examples of employer-based programs are monetary compensation for the bicycle (which is higher than for the car), providing public transport passes and a bicycle plan (providing bicycles). Moreover, after Covid-19, more and more companies are reconsidering their policy on limiting the use of company cars (NOS, 2021).

By applying these kinds of policies by employers, multimodal behaviour could also be enhanced, and it can create a stimulus for travellers away from being in a more car-dependant travel user class. It is suggested to employ the mentioned programs as employers. For governments, it is suggested to improve employers' willingness to do so, especially among the smaller companies, which are likely to have a (lack) of available mode alternatives for their employees.

After all, specifically targeted policies can be set up based on our policy directions, which is essential as not only one average group is analysed, but we found travel users classes which characterise several groups in society with different needs and preferences. Still, it should be noted that due to the probabilistic nature of the used latent class analysis and having average and grouped classes, some individuals are likely to fit in none or multiple identified classes, which are more challenging to target. Moreover, the effectiveness of specific interventions for specific classes is unspecified yet. This could be addressed in further research before applying policies to make multimodal behaviour less exclusive, as there is an unequal distributed multimodal trend among different socio-demographic groups in society.

6.3 Discussion

First, our results were discussed (section 6.1), potentially in light of previous work. This study emphasises several important findings in the research of multimodal travel behaviour, explained in this chapter. Nonetheless, this study has some inherent limitations; some have little impact, and others give room for future research or new research aims. The following reflection aspects are made: (measured) multimodal travel behaviour, data and sample, the definition of travel behaviour, determinants of multimodal travel behaviour and trends.

(Measured) multimodal travel behaviour

Multimodality (measurements) provides information about the diversity (of mode usages) and/or equality (balanced mode use). Choices about which measure (from which measurement category) to use were made based on our specific research field and the method used (Chapter 2). Our study showed that using several (different) measurements can overcome their limitations alone, and using latent class analysis can provide more information on which mode intensities the measure values are based on. Moreover, both used multimodality measures showed reasonably similar results. Some other aspects of defining multimodality and measuring it remain to be explored.

First, to guide in the (combinational) use of several measures, a refined decision-based scheme about which measures to use to define multimodality based on the research field, situation, data aspects, other methodologies, or other requirements could be set up by researchers with an overview of the (dis)advantages and suitability aspects.

Second, it is not yet indicative in the multimodality definition/measuring if some modes are more desirable, while, on the other hand, some modes are mostly not included. Now only attention is paid to the diversity of mode usage and equality of mode intensities (like most research). However, it can be argued that (unimodal) active travelling (walking and cycling) is considered more desirable because of sustainability, social, health-related and economic aspects, compared to car use, but also to public transport use, which is not well reflected in the multimodality measure. But, our research did effectively include the travel patterns behind the measurements. Moreover, our analysis used the main travel modes for measuring multimodality. It excluded individuals with their full travel patterns if it included other or relatively uncommon modes like inline skates (active travel) or (public transportation) boats. Because the multimodality measurement is affected by using many modes or how modes are grouped, as well as less clear classes would emerge. Other upcoming modes, like electric cars or car-sharing services, are generally regarded as more sustainable, affecting the notion of multimodality. How to deal with individuals using more desirable or relatively uncommon modes in multimodal measuring remains to be determined and gives room for further research.

Data and sample

Our study used data from the Dutch National Travel Survey (OViN) of 2010-2017, and (dis) advantages of this dataset are already pointed out in section 4.3. Some of the implications of using this data are also reflected upon in the upcoming paragraphs, besides two aspects of the data (sample and representativeness) which are now elaborated.

First, regarding the used sample for our data, it is already reflected upon that some studies incorporate participants from adolescents onwards, and others only incorporate adults, like we did, as children likely have different travel patterns compared to adults. Moreover, when having included both children and adults, classes comprising the same modal patterns are preferably also targeted specifically for children and adults, next by targeting the different classes differently, which is not a straightforward task when adults and children are in the same class. Furthermore, it is likely that in the life phase, from primary education to secondary education and from secondary education to continuation education (potentially with a driving license), children could transition towards new travel patterns. Hence, analysing children (or adolescents) as one group (mostly done) is not representative.

When the aim is to analyse the multimodal travel patterns of children of several ages more in-depth, in a study comparable to ours, a point of departure could be by including information on the household level for this new research aim. As OViN is an individual household survey from 2010-2017, the research set up from before, MON 2004-2009 (in Dutch: Mobiliteitsonderzoek Nederland), could be used as this comprises travel diaries for all household members (Statistics Netherlands, 2018), although being less recent and not comparable to the OViN. It could thus be accounted for that the travel behaviour of adults is influenced by travelling together with children to activities by including a descriptive trip purpose variable based on having travelled together for children's activities, for example.

From another viewpoint, diaries for all household members could help analyse multimodal travel patterns of several groups (including children and adults) because parents' resources, preferences and habits are likely to influence the children's travel patterns, also when becoming an adult. So, including a broader set of variables in questionnaires, measuring the attitudinal comparability of households towards modes, for instance, could provide more information on the influence of parent-child relations in travel patterns. A study, with another context and method investigated the influence of parental attitudes and the multimodal travel patterns of children to school (Mehdizadeh & Ermagun, 2020). By applying a new perspective in research, similar to ours but with another aim, the attitudinal theories could enhance the understanding of travel behaviour.

Second, as explained in section 4.8, our sample represents, in general, the whole population, but no entirely accurate view of this could be acquired. Potential underrepresentedness (and others being overrepresented) might be due to selective non-responses, as some groups are less inclined to fill such travel diaries when selected (Statistics Netherlands, 2018). Still, this high-quality survey, with many respondents, gives one of the highest possible representative results compared to other studies using other surveys or questionnaires. This non-fully representativeness implies that our results are less able to be generalised to the total Dutch population, and some results might not reflect the travel patterns well of specific individuals with certain characteristics.

For future research similar to ours, an improvement possibility could be to use the population weights based on personal characteristics provided in OViN (Statistics Netherlands, 2018). By setting the sample weight option on in the latent class estimation (Vermunt & Magidson, 2016), it could be accounted for that some individuals are taken more often into account in the analysis when they are underrepresented our dataset used and the other way around, although the weights cannot be entirely regarded as accurate, to acquire a view of the broader population.

Definition of travel behaviour

The used operationalisation of travel behaviour, with corresponding choices and scoping (section 4.4), has some implications, where the main points are explained here, and some other aspects briefly come through in upcoming paragraphs.

As explained, this research focussed on daily travel in The Netherlands, like commuting or recreational travel. Travel users doing daily activities (just outside) the border, which is foreign travel, are thus not included, and no specific policy directions can be created for those people based on our results. These people are mainly individuals living closer to the border of The Netherlands (with Germany and Belgium). They are likely less represented in the data and are not included to avoid having biased results. However, it is likely that when living in those areas, most people are more habitual differently and tend to travel (probably by car in these less urban dense areas) abroad for groceries or refuelling, for instance, as in general several activities are cheaper 'abroad'.

In order to target these individuals specifically, it is suggested to analyse the multimodal travel patterns separately with a new research aim with another or new survey by researchers or institutes to assure more representativeness of these (groups of) people. Some research has already been done about cross-border work-purpose travel to The Netherlands (Statistics Netherlands, 2020b). By analysing the mentioned individuals, doing cross-border travel from The Netherlands for several daily purposes, separately, it can be accounted for the previously mentioned by assessing more contextual-specific determinants of multimodal travel behaviour for multiple purposes.

Our study analysed the travel behaviour on the stage level, meaning that all modes are taken into account from all trips and parts of trips (stages) made by individuals. For instance, a trip can comprise a cycling stage, public transport stage and walking stage, compared to someone using only the car (one stage) in a trip. In this way, as earlier explained and proposed by researchers, the full variability in travel patterns (multimodality) can be captured instead of only looking at the main modes used for every trip. An implication, as pointed out by a study which used the trip level (e.g., De Haas et al., 2018), is that less comprehensive information is provided by choosing this operationalising. Because the number of stages is larger in multimodal behaviour compared to unimodal behaviour, the stage level does not reflect the number of trips (showing the behaviour of someone) made in both patterns. Additionally, at the stage level, it is unclear which mode is used as access, egress or main mode in a class comprising modal patterns. Another point of consideration we can think of is that our study did not show the purposes of the stages, which is in OViN measured on the trip level. Other researchers showed the occurrence of purposes per trip in several classes for working, shopping, leisure, and other trips (e.g., De Haas et al., 2018).

Especially when analysing travel behaviour at the stage level and the willingness to provide more information about the patterns, an opportunity lies in giving more descriptive statistics per class in such a latent class analysis as we did. For instance, about the number of trips made per mode (next to the number of stages like we did); the percentage of how often this mode is used as a main mode in a trip; the percentage showing how many trips were multimodal (involving two or more modes); and the percentage showing the trip purpose distribution, to give a broader view on multimodal behaviour and policies can be even more targeted.

Several choices were made and explained regarding the measure of the intensity of travel behaviour and the mode set which is analysed. The number of stages of a mode is used instead of the travel distance or travel times per mode, as the self-reported nature of travel distances and travel times leads to more flawed results. The modes set analysed is walk, bicycle, public transport (bus, tram, metro and train) and car (passenger and driver), as more often done in other studies. However, other research involving travel user latent classes used travel distances to additionally profile classes (e.g., Ton et al., 2020) or divided public transport into bus/tram/metro (mainly inter-city travel) and train (mainly intra-city travel) (e.g., Molin et al., 2016), as public transport usually has lower and higher distances and combining them gives an overrepresentation in terms of distance. This implies that our results are less able, to some extent, to have a contextualised understanding of travel behavioural patterns.

A window remains for using travel distances as a descriptive variable in latent class analysis, like our research, to better profile the identified classes by knowing if travel users travel more within a city or between cities, which can be targeted differently. A deeper understanding of what kind of distances are travelled with what modes can thus be acquired. For instance, a pattern comprising only short bicycle trips is not per se suited to switch to more multimodal behaviour. On the other hand, long-distance car trips are likely harder to switch to more multimodal options, while this pattern has more room for improvement.

Determinants of multimodal travel behaviour

Our study has set up a widely used set of determinants and correlates with multimodality (Chapter 3 and Chapter 4). However, when using latent class analysis, it is unclear if all potential variables are used, but the risk of overfitting with too many variables should not be forgotten. Nevertheless, our findings suggest that mobility resource availability plays a significant role in the prediction of multimodality.

The ones included are licensure, ownership of a household car/company car/household bicycle. However, unlike other studies, not all relevant mobility resource data is included. For instance, only data for the ownership and kind of student public transport pass was available (and is not included), while others included the ownership of (kinds of) public transport pass or subscription in general in the analysis (e.g., An et al., 2022; Heinen & Chatterjee, 2015; Ton et al., 2020). Moreover, other mobility resource-related influences which are not part of our data, such as distances to train stations (can also be regarded as built-environment characteristics) (De Haas et al., 2018) or (multimodal) cycling infrastructural budgets (not fully about mobility resources, but it is about the resource of infrastructure) of municipalities in regions of travellers, could likely affect multimodal behaviour. Before considering or identifying possibilities in adding more variables to the model, as touched upon earlier, the notion of causation should be explained.

Buehler and Hamre (2015) state that the direction of causation between the explanatory variables and the dependant (latent class, in our case) variable is not fully clear. Due to self-selection, travellers who wish to use public transport might decide to live in dense urban areas, as the notion of residential self-selection is also explained by Molin et al. (2016) and Faber et al. (2021). The self-selection and endogeneity, as explained by Klinger (2017), are shown in another example of mobility resource ownership. Someone might have certain attitudes towards multimodal behaviour and can be (un)likelier to purchase a bicycle. In other words, due to self-selection, someone that likes to have a specific mobility style is likely to acquire the several modes to be used (the other way around than assumed and specified in the model). So, owning several mobility options cannot be regarded entirely as fully exogenous in determining the latent class with respecting modal patterns.

As part of our model, we included urban density and mobility resources in our analyses. This is less of an issue for urban density, as many researchers include it as part of the model and assume it to be exogenous (e.g., De Haas et al., 2018; Ton et al., 2020). However, in other latent class models, mobility resource variables are included as descriptive only, as in this way, LCA can overcome the issue by not regarding it as exogenous in the model (e.g., Kroesen, 2014). Some others included it in their modes, in clustering analyses (e.g., An et al., 2022) or regression analyses (e.g., An et al.; Scheiner et al., 2016). However, no possibilities of handling causation issues exist for those methods. Room for improvement lies in testing our model to see how the results change when the mobility resources were not included in the model (regarded as endogenous and not exogenous) and added as descriptives to profile the classes.

Besides the issue of self-selection, it can be derived that attitudinal factors could play a role in determining multimodal travel behaviour. Potential physiological factors are travel-related location reasons (i.e., the extent to which a travel preference affect the decision to move to a location) and attitudes towards modes (e.g., Faber et al., 2021), which could determine mode usage. Other examples are seen, for instance, in other research which has been done about mode attitudes concerning multimodal travel patterns (e.g., Molin et al., 2016; Ton et al., 2020). It could be assessed if people are likely to travel with their most preferred mode, as research suggests that this is not always the case, which is called cognitive dissonance (An et al., 2021, 2022). On the other hand, Ton et al. (2020) show that mode consonance (ideal match between attitudes and behaviour) varies among their founded classes involving multimodal travel patterns. Including attitudinal and psychological factors, a substantive other research direction could complement our identified multimodal travel patterns. However, no data is available for this as perceptions and preferences are not included in OViN, to-be-made questionnaires or the Mobility Panel Netherlands could be used.

Trends

By having identified aggregate modality styles in our study, more generic population groups are captured and represent 2010-2017. Specifically, the estimates are based on the pooled years of cross-sectional data and provide information about the travel patterns and determinants once (statically) at the population level. Because of the inclusion of the year variable in latent class analysis, we were enabled to show how the classes involving multimodal travel patterns developed over time in terms of class sizes. Using a large and representative (other) population set allowed us to comprehensively view multimodal trends across society. It provides a general trace of how class sizes develop over time. Although the aggregate modality styles can be intuitively interpreted, other aspects show a more limited view on the individual level, as some are already explained in section 5.5.

The trends analysis could be continued by analysing the growth of class sizes on the individual level by employing or acquiring longitudinal panel data, which traces individuals for a more extended period. For instance, a Latent Class growth model could be employed, ideally suited to deal with longitudinal data where time-specific latent classes can be acquired (Vermunt & Magidson, 2016). This way, the trend overview could be advanced by capturing individual trends where class membership involving multimodal travel patterns depends on time.

Literature

An, Z., Heinen, E., & Watling, D. (2021). When you are born matters: An age-period-cohort analysis of multimodality. *Travel Behaviour and Society*, *22*, 129-145. https://doi.org/10.1016/j.tbs.2020.09.002

An, Z., Heinen, E., & Watling, D. (2022). Multimodal travel behaviour, attitudes, and cognitive dissonance. *Transportation Research Part F: Traffic Psychology and Behaviour*, *91*, 260-273. https://doi.org/10.1016/j.trf.2022.10.0

An, Z., Heinen, E., & Watling, D. (2023). The level and determinants of multimodal travel behavior: Does trip purpose make a difference? *International Journal of Sustainable Transportation*, *17*(2), 103-117. https://doi.org/10.1080/15568318.2021.198519

Anable, J. (2005). 'Complacent Car Addicts' or 'Aspiring Environmentalists'? Identifying travel behaviour segments using attitude theory. *Transport Policy*, *12*(1), 65-78. https://doi.org/10.1016/j.tranpol.2004.11.004

Araghi, Y., Kroesen, M., & van Wee, B. (2017). Identifying reasons for historic car ownership and use and policy implications: An explorative latent class analysis. *Transport Policy*, *56*, 12-18. https://doi.org/10.1016/j.tranpol.2017.02.008

Arntzen, K., Hoekstra, A., Jansen, H., Linden, K. van der, Tromp, A., Verhoeven, T., & Verroen, E. (2020). *De multimodale Hub en Rijkswaterstaat*. Rijkswaterstaat. https://open.rijkswaterstaat.nl/publish/pages/163586/rws_multimodale_hub_juni_2020_mr_090620.pdf

Brand, C., Goodman, A., Rutter, H., Song, Y., & Ogilvie, D. (2013). Associations of individual, household and environmental characteristics with carbon dioxide emissions from motorised passenger travel. *Applied Energy*, *104*, 158-169. https://doi.org/10.1016/j.apenergy.2012.11.001

Buehler, R., & Hamre, A. (2015). The multimodal majority? Driving, walking, cycling, and public transportation use among American adults. *Transportation*, *42*(6), 1081-1101. https://doi.org/10.1007/s11116-014-9556-z

Buehler, R., & Pucher, J. (2012). Walking and Cycling in Western Europe and the United States. *TR News*, *280*, 34-42.

Chatterjee, S., & Hadi, A.S. (1991). *Regression Analysis by Example* (5th ed.). Wiley.

Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3rd ed.). Sage Publications.

Curtis, C., & Headicar, P. (1997). Targeting travel awareness campaigns. *Transport Policy*, *4*(1), 57-65. https://doi.org/10.1016/S0967-070X(96)00034-0

De Haas, M. C., Kroesen, M., Chorus, C., Hoogendoorn-Lanser, S., & Hoogendoorn, S. (2022). E-bike user groups and substitution effects: Evidence from longitudinal travel data in the Netherlands. *Transportation*, *49*(3), 815-840. https://doi.org/10.1007/s11116-021-10195-3

De Haas, M. C., Scheepers, C. E., Harms, L. W. J., & Kroesen, M. (2018). Travel pattern transitions: Applying latent transition analysis within the mobility biographies framework. *Transportation Research Part A: Policy and Practice*, *107*, 140-151. https://doi.org/10.1016/j.tra.2017.11.007

Diana, M., & Mokhtarian, P. L. (2008). Travelers' segmentation based on multimodality behaviors and attitudes. University of California, Institute of Transportation Studies. https://escholarship.org/uc/item/2cb1z6v7

Diana, M., & Mokhtarian, P. L. (2009). Desire to change one's multimodality and its relationship to the use of different transport means. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(2), 107-119. https://doi.org/10.1016/j.trf.2008.09.001

Diana, M., & Pirra, M. (2016). A comparative assessment of synthetic indices to measure multimodality behaviours. *Transportmetrica A: Transport Science*, *12*(9), 771-793. https://doi.org/10.1080/23249935.2016.1177133

ECF. (2018). *The benefits of cycling*. https://ecf.com/sites/ecf.com/files/TheBenefitsOfCycling_final-v2.pdf

EEA. (n.d.). *Greenhouse gas emissions from transport in Europe*. Retrieved February 27, 2023, from https://www.eea.europa.eu/ims/greenhouse-gas-emissions-from-transport

Elghozi, T. (2021, June 30). *Tracking efficiency in the transport sector* [slides]. IEA. https://iea.blob.core.windows.net/assets/6ebb6287-fdd5-4934-850b-7db94d89ad94/2B.1_IEA_Trackingefficiencyinthetransportsector.pdf

European Commission. (n.d.). *2018 - Year of Multimodality*. Retrieved March 1, 2023, from https://transport.ec.europa.eu/transport-themes/logistics-and-multimodal-transport/2018-year-multimodality_en

European Parliament. (2023, February 16). *MEPs say cycling will help the EU's green transition.* https://www.europarl.europa.eu/news/en/press-room/20230210IPR74808/meps-say-cycling-will-help-the-eu-s-green-transition

Faber, R., Merkies, R., Damen, W., Oirbans, L., Massa, D., Kroesen, M., & Molin, E. (2021). The role of travel-related reasons for location choice in residential self-selection. *Travel Behaviour and Society*, *25*, 120-132. https://doi.org/10.1016/j.tbs.2021.07.003

Faber, R. M., Jonkeren, O., de Haas, M. C., Molin, E. J. E., & Kroesen, M. (2022). Inferring modality styles by revealing mode choice heterogeneity in response to weather conditions. *Transportation Research Part A: Policy and Practice*, *162*, 282-295. https://doi.org/10.1016/j.tra.2022.06.003

Gerlofs-Nijland, M., Staatsen, B., Geelen, L., Mulder, Y., Fuse, V., Sharashidze, N., Bonvoisin, N., Choi, A. K. M., & Fruianu, M. (2021). *Road transport facts and figures. How healthy and environmentally friendly is our transport today?* RIVM. https://thepep.unece.org/sites/default/files/2021-05/eMagazine%20Road%20transport%20facts%20%20figures_updated%2017%20May%202021_1.pdf

Götschi, T., de Nazelle, A., Brand, C., Gerike, R., & PASTA Consortium. (2017). Towards a Comprehensive Conceptual Framework of Active Travel Behavior: A Review and Synthesis of Published Frameworks. *Current Environmental Health Reports*, *4*(3), 286-295. https://doi.org/10.1007/s40572-017-0149-9

Haustein, S., & Kroesen, M. (2022). Shifting to more sustainable mobility styles: A latent transition approach. *Journal of Transport Geography*, *103*, 103394. https://doi.org/10.1016/j.jtrangeo.2022.103394

Heinen, E. (2018). Are multimodals more likely to change their travel behaviour? A cross-sectional analysis to explore the theoretical link between multimodality and the intention to change mode choice. *Transportation Research Part F: Traffic Psychology and Behaviour*, *56*, 200-214. https://doi.org/10.1016/j.trf.2018.04.010

Heinen, E., & Chatterjee, K. (2015). The same mode again? An exploration of mode choice variability in Great Britain using the National Travel Survey. *Transportation Research Part A: Policy and Practice*, *78*, 266-282. https://doi.org/10.1016/j.tra.2015.05.015

Heinen, E., & Mattioli, G. (2019a). Does a high level of multimodality mean less car use? An exploration of multimodality trends in England. *Transportation*, *46*(4), 1093-1126. https://doi.org/10.1007/s11116-017-9810-2

Heinen, E., & Mattioli, G. (2019b). Multimodality and CO2 emissions: A relationship moderated by distance. *Transportation Research Part D: Transport and Environment*, *75*, 179-196. https://doi.org/10.1016/j.trd.2019.08.022

Heinen, E., & Ogilvie, D. (2016). Variability in baseline travel behaviour as a predictor of changes in commuting by active travel, car and public transport: A natural experimental study. *Journal of Transport & Health*, *3*(1), 77-85. https://doi.org/10.1016/j.jth.2015.11.002

Hunecke, M., Groth, S., & Wittowsky, D. (2020). Young social milieus and multimodality: Interrelations of travel behaviours and psychographic characteristics. *Mobilities*, *15*(3), 397-415. https://doi.org/10.1080/17450101.2020.1732099

IEA. (n.d.). *Transport - Improving the sustainability of passenger and freight transport*. Retrieved February 26, 2023, from https://www.iea.org/topics/transport

IEA. (2022, September). *Transport - Analysis*. Tracking report. https://www.iea.org/reports/transport

IPCC. (2014). Transport. In O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel & J. C. Minx (Eds.), *Climate Change 2014: Mitigation of Climate Change*. (pp. 599-670). Cambridge University Press.

IPCC. (2022). Summary for Policymakers. In H.-O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, & B. Rama (Eds.), *Climate Change 2022: Impacts, Adaptation and Vulnerability*. (pp. 3-33). Cambridge University Press.

Kennisinstituut voor Mobiliteitsbeleid. (2019). *Mobiliteitsbeeld 2019*. https://www.kimnet.nl/binaries/kimnet/documenten/rapporten/2019/11/12/mobiliteitsbeeld-2019-vooral-het-gebruik-van-de-trein-neemt-toe/Mobiliteitsbeeld+2019.pdf

Klinger, T. (2017). Moving from monomodality to multimodality? Changes in mode choice of new residents. *Transportation Research Part A: Policy and Practice*, *104*, 221-237. https://doi.org/10.1016/j.tra.2017.01.008

Klingert, L. (2022, March 16). *UN adopts resolution to encourage cycling to combat global warming*. The Brussels Times. https://www.brusselstimes.com/210961/un-adopts-resolution-to-encourage-cycling-to-combat-global-warming

Kroesen, M. (2014). Modeling the behavioral determinants of travel behavior: An application of latent transition analysis. *Transportation Research Part A: Policy and Practice*, *65*, 56-67. https://doi.org/10.1016/j.tra.2014.04.010

Kroesen, M. (2019). Is active travel part of a healthy lifestyle? Results from a latent class analysis. *Journal of Transport & Health*, *12*, 42-49. https://doi.org/10.1016/j.jth.2018.11.006

Kroesen, M., & van Cranenburgh, S. (2016). Revealing transition patterns between mono- and multimodal travel patterns over time: A mover-stayer model. *European Journal of Transport and Infrastructure Research, 16*(4). https://doi.org/10.18757/ejtir.2016.16.4.3169

Kroesen, M., & van Wee, B. (2022). Understanding how accessibility influences health via active travel: Results from a structural equation model. *Journal of Transport Geography*, *102*, 103379. https://doi.org/10.1016/j.jtrangeo.2022.103379

Kuhnimhof, T., Armoogum, J., Buehler, R., Dargay, J., Denstadli, J. M., & Yamamoto, T. (2012). Men Shape a Downward Trend in Car Use among Young Adults-Evidence from Six Industrialized Countries. *Transport Reviews*, *32*(6), 761-779. https://doi.org/10.1080/01441647.2012.736426

Kuhnimhof, T., Buehler, R., & Dargay, J. (2011). A New Generation: Travel Trends for Young Germans and Britons. *Transportation Research Record*, *2230*(1), 58-67. https://doi.org/10.3141/2230-07

Kuhnimhof, T., Buehler, R., Wirtz, M., & Kalinowska, D. (2012). Travel trends among young adults in Germany: Increasing multimodality and declining car use for men. *Journal of Transport Geography*, *24*, 443-450. https://doi.org/10.1016/j.jtrangeo.2012.04.018

Kuhnimhof, T., Chlond, B., & Huang, P.-C. (2010). Multimodal Travel Choices of Bicyclists. Multiday Data Analysis of Bicycle Use in Germany. *Transportation Research Record*, *2190(1)*, 19-27. https://doi.org/10.3141/2190-03

Lee, S. (2022). Exploring Associations between Multimodality and Built Environment Characteristics in the U.S. *Sustainability, 14*(11), 6629. https://doi.org/10.3390/su14116629

Lee, Y., Circella, G., Mokhtarian, P. L., & Guhathakurta, S. (2020). Are millennials more multimodal? A latent-class cluster analysis with attitudes and preferences among millennial and Generation X commuters in California. *Transportation*, *47*(5), 2505-2528. https://doi.org/10.1007/s11116-019-10026-6

Lemonde, C., Arsenio, E., & Henriques, R. (2021a). Integrative analysis of multimodal traffic data: Addressing open challenges using big data analytics in the city of Lisbon. *European Transport Research Review*, *13*(1), 64. https://doi.org/10.1186/s12544-021-00520-3

Lemonde, C., Arsenio, E., & Henriques, R. (2021b). Public Transportation Multimodality in the City of Lisbon. *Transportation Research Procedia*, *58*, 75-82. https://doi.org/10.1016/j.trpro.2021.11.011

Lezhnina, O., & Kismihók, G. (2022). Latent Class Cluster Analysis: Selecting the number of clusters. *MethodsX*, *9*, 101747. https://doi.org/10.1016/j.mex.2022.101747

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., AlMazroa, M. A., Amann, M., Anderson, H. R., Andrews, K. G., Aryee, M., Atkinson, C., Bacchus, L. J., Bahalim, A. N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M. L., … Ezzati, M. (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: A systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, *380*(9859), 2224-2260. https://doi.org/10.1016/S0140-6736(12)61766-8

Macharis, C., & De Witte, A. (2012). The typical company car user does not exist: The case of Flemish company car drivers. *Transport Policy*, *24*, 91-98. https://doi.org/10.1016/j.tranpol.2012.06.018

Magidson, J., & Vermunt, J. (2002). Latent class models for clustering: A comparison with K-means. *Canadian Journal of Marketing Research, 20*(1), 36-43.

Magidson, J., & Vermunt, J. K. (2004). Latent class models. In D. Kaplan (Ed.), *The Sage Handbook of Quantitative Methodology for the Social Sciences* (pp. 175-198). Sage.

Maibach, E., Steg, L., & Anable, J. (2009). Promoting physical activity and reducing climate change: Opportunities to replace short car trips with active transportation. *Preventive Medicine*, *49*(4), 326-327. https://doi.org/10.1016/j.ypmed.2009.06.028

Mehdizadeh, M., & Ermagun, A. (2020). "I'll never stop driving my child to school": On multimodal and monomodal car users. *Transportation*, *47*(3), 1071-1102. https://doi.org/10.1007/s11116-018-9949-5

Ministerie van Infrastructuur en Waterstaat. (2019a). *Mobility as a Service (MaaS): multimodaal reisadvies op maat*. Rijksoverheid. https://www.rijksoverheid.nl/onderwerpen/mobiliteit-nu-en-in-de-toekomst/mobility-as-a-service-maas

Ministerie van Infrastructuur en Waterstaat. (2019b). *MaaS-pilots - Optimaliseren van het mobiliteitssysteem*. https://open.overheid.nl/repository/ronl-747e2350-a2fe-4048-ab91-63bfef856e87/1/pdf/Brochure%20MaaS-pilots%20NL%20190522.pdf

Molin, E., Mokhtarian, P., & Kroesen, M. (2016). Multimodal travel groups and attitudes: A latent class cluster analysis of Dutch travelers. *Transportation Research Part A: Policy and Practice*, *83*, 14-29. https://doi.org/10.1016/j.tra.2015.11.001

Nobis, C. (2007). Multimodality: Facets and Causes of Sustainable Mobility Behavior. *Transportation Research Record: Journal of the Transportation Research Board*, *2010*(1), 35-44. https://doi.org/10.3141/2010-05

NOS. (2021, May 17). *Corona raakt ook de auto van de zaak*. https://nos.nl/nieuwsuur/artikel/2381071-corona-raakt-ook-de-auto-van-de-zaak

NOS. (2022, April 19). *Grote werkgevers moeten personeel milieuvriendelijker laten reizen*. https://nos.nl/artikel/2425727-grote-werkgevers-moeten-personeel-milieuvriendelijker-laten-reizen

NOS. (2023, June 22). *Werknemer moet de auto uit: Kilometervergoeding fietsers steeds hoger*. https://nos.nl/artikel/2479826-werknemer-moet-de-auto-uit-kilometervergoeding-fietsers-steeds-hoger

NS. (2021). *Vervoersplan NS 2022*. https://www.ns.nl/binaries/_ht_1641202755768/content/assets/ns-nl/over-ons/ns-vervoerplan-2022.pdf

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(4), 535-569. https://doi.org/10.1080/10705510701575396

Oberski, D. (2016). Mixture Models: Latent Profile and Latent Class Analysis. In *Modern Statistical Methods for HCI* (pp. 275-287). Springer International Publishing. http://link.springer.com/10.1007/978-3-319-26633-6_12

Oja, P., Titze, S., Bauman, A., de Geus, B., Krenn, P., Reger-Nash, B., & Kohlberger, T. (2011). Health benefits of cycling: A systematic review. *Scandinavian Journal of Medicine & Science in Sports*, *21*(4), 496-509. https://doi.org/10.1111/j.1600-0838.2011.01299.x

Olafsson, A. S., Nielsen, T. S., & Carstensen, T. A. (2016). Cycling in multimodal transport behaviours: Exploring modality styles in the Danish population. *Journal of Transport Geography*, *52*, 123-130. https://doi.org/10.1016/j.jtrangeo.2016.03.010

Olde Kalter, M.-J., La Paix Puello, L., & Geurs, K. T. (2020). Do changes in travellers' attitudes towards car use and ownership over time affect travel mode choice? A latent transition approach in the Netherlands. *Transportation Research Part A: Policy and Practice*, *132*, 1-17. https://doi.org/10.1016/j.tra.2019.10.015

Prato, C. G., Halldórsdóttir, K., & Nielsen, O. A. (2017). Latent lifestyle and mode choice decisions when travelling short distances. *Transportation*, *44*(6), 1343-1363. https://doi.org/10.1007/s11116-016-9703-9

Pucher, J., Dill, J., & Handy, S. (2010). Infrastructure, programs, and policies to increase bicycling: An international review. *Preventive Medicine*, *50*, S106-S125. https://doi.org/10.1016/j.ypmed.2009.07.028

Ren, Z., Fusco, G., Lownes, N., & Zhu, J. (2022). Entropy-Based Diversity Quantification of Multimodal Transportation Systems: Physical Infrastructure Perspective versus Travel Behavior Perspective. *Journal of Urban Planning and Development*, *148*(3), 04022029. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000855

Rijkswaterstaat Duurzame Mobiliteit. (n.d.). *Multimodale hubs*. Retrieved July 13, 2023, from https://rwsduurzamemobiliteit.nl/praktijk/veranderen/multimodale-hubs/

S., L. (2023). a blurry photo of a city street with a bus [Photograph]. https://unsplash.com/photos/VRU7TWK11Cg. Licensed under Unsplash License.

Saunders, L. E., Green, J. M., Petticrew, M. P., Steinbach, R., & Roberts, H. (2013). What are the health benefits of active travel? A systematic review of trials and cohort studies. *PLoS One, 8*(8), e69912. https://doi.org/10.1371/journal.pone.0069912

Schafer, A., & Victor, D. G. (1999). Global passenger travel: Implications for carbon dioxide emissions. *Energy, 24*(8), 657-679. https://doi.org/10.1016/S0360-5442(99)00019-5

Scheiner, J., Chatterjee, K., & Heinen, E. (2016). Key events and multimodality: A life course approach. *Transportation Research Part A: Policy and Practice*, *91*, 148-165. https://doi.org/10.1016/

Schlich, R., & Axhausen, K. W. (2003). *Habitual travel behaviour: Evidence from a six-week travel diary*. *Transportation, 30*(13-36). https://doi.org/10.1023/A:1021230507071

Schneider, F., Ton, D., Zomer, L.-B., Daamen, W., Duives, D., Hoogendoorn-Lanser, S., & Hoogendoorn, S. (2021). Trip chain complexity: A comparison among latent classes of daily mobility patterns. *Transportation*, *48*(2), 953-975. https://doi.org/10.1007/s11116-020-10084-1

Shukla, S., & Naganna, S. (2014). A Review ON K-means DATA Clustering APPROACH. *International Journal of Information & Computation Technology, 4*(17), 1847-1860. ISSN 0974-2239.

Sinha, P., Calfee, C. S., & Delucchi, K. L. (2021). Practitioner's Guide to Latent Class Analysis: Methodological Considerations and Common Pitfalls. *Critical Care Medicine*, *49*(1), e63-e79. https://doi.org/10.1097/CCM.0000000000004710

Statistics Netherlands. (2014). *Addendum betreffende fietsbezit OViN 2013 en 2014.* DANS. https://doi.org/10.17026/dans-x95-5p7y

Statistics Netherlands. (2018). *Onderzoek Verplaatsingen in Nederland 2017: Onderzoeksbeschrijving*. DANS. https://doi.org/10.17026/dans-xxt-9d28

Statistics Netherlands. (2020a, March 12). *Car fleet growing faster than population.* https://www.cbs.nl/en-gb/news/2020/10/car-fleet-growing-faster-than-population

Statistics Netherlands. (2020b, November 18). *Inbound commuters often live just across the border.* https://www.cbs.nl/en-gb/news/2020/46/inbound-commuters-often-live-just-across-the-border

Statistics Netherlands & Department of Waterways and Public Works. (2010*). Netherlands Travel Survey (OViN) 2010 version 2.0* [Dataset]. DANS. https://doi.org/10.17026/dans-zhs-ghwg

Statistics Netherlands & Department of Waterways and Public Works. (2011). *Netherlands Travel* Survey (OViN) *2011 version 2.0* [Dataset]. DANS. https://doi.org/10.17026/dans-xv2-hapb

Statistics Netherlands & Department of Waterways and Public Works. (2012). *Netherlands Travel Survey (OViN) 2012 version 2.0* [Dataset]. DANS. https://doi.org/10.17026/dans-2bs-q7u2

Statistics Netherlands & Department of Waterways and Public Works. (2013). *Netherlands Travel Survey (OViN) 2013* [Dataset]. DANS. https://doi.org/10.17026/dans-x9h-dsdg

Statistics Netherlands & Department of Waterways and Public Works. (2014). *Netherlands Travel Survey (OViN) 2014* [Dataset]. DANS. https://doi.org/10.17026/dans-x95-5p7y

Statistics Netherlands & Department of Waterways and Public Works. (2015). *Netherlands Travel Survey (OViN) 2015 version 2.0* [Dataset]. DANS. https://doi.org/10.17026/dans-z2v-c39p

Statistics Netherlands & Department of Waterways and Public Works. (2016). *Netherlands Travel Survey (OViN) 2016* [Dataset]. DANS. https://doi.org/10.17026/dans-293-wvf7

Statistics Netherlands & Department of Waterways and Public Works. (2017). *Netherlands Travel Survey (OViN) 2017* [Dataset]. DANS. https://doi.org/10.17026/dans-xxt-9d28

Streit, T., Allier, C.-E., Weiss, C., Chlond, B., & Vortisch, P. (2015). Changes in Variability and Flexibility of Individual Travel in Germany: Trends and Drivers. *Transportation Research Record*, *2496*(1), 10-19. https://doi.org/10.3141/2496-02

Susilo, Y. O., & Axhausen, K. W. (2014). Repetitions in individual daily activity-travel-location patterns: A study using the Herfindahl-Hirschman Index. *Transportation*, *41*(5), 995-1011. https://doi.org/10.1007/s11116-014-9519-4

The Netherlands' Cadastre, Land Registry and Mapping Agency & Statistics Netherlands. (2019). *Wijk- en buurtkaart 2017* [Dataset]. Statistics Netherlands. https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/wijk-en-buurtkaart-2017

Ton, D., Lara-Britt, Z., Schneider, F., Sascha, H.-L., Dorine, D., Oded, C., & Hoogendoorn, S. (2020). Latent classes of daily mobility patterns: The relationship with attitudes towards modes. *Transportation*, *47*(4), 1843-1866. https://doi.org/10.1007/s11116-019-09975-9

UNEP. (2019, June 19). *Cycling, the better mode of transport*. http://www.unep.org/news-and-stories/story/cycling-better-mode-transport

United Nations Economic Commission for Europe. (2021). *Recommendations for Green and Healthy Sustainable Transport - "Building Forward Better"*. United Nations. https://doi.org/10.18356/9789210056915

Van Eenoo, E., Boussauw, K., & Fransen, K. (2022). Car dependency beyond land use: Can a standardized built environment indicator predict car use? *Journal of Transport and Land Use*, *15*(1), 117-136. https://doi.org/10.5198/jtlu.2022.2073

Van Eenoo, E., Fransen, K., & Boussauw, K. (2022). Perceived car dependence and multimodality in urban areas in Flanders (Belgium). *European Journal of Transport and Infrastructure Research*, *22*(1), 42-62. https://doi.org/10.18757/EJTIR.2022.22.1.6044

Van Wee, B., & Banister, D. (2016). How to Write a Literature Review Paper? *Transport Reviews*, *36*(2), 278-288. https://doi.org/10.1080/01441647.2015.1065456

Vermunt, J. K. (2004). Toepassingen van latent klasse analyse in sociaal wetenschappelijk onderzoek. *Sociale Wetenschappen, 47,* 2-14.

Vermunt, J. K. (2010). Latent Class Modeling with Covariates: Two Improved Three-Step Approaches. *Political Analysis*, *18*(4), 450-469. https://doi.org/10.1093/pan/mpq025

Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. In J. Hagenaars, & A. McCutcheon (Eds.), *Applied latent class analysis* (pp. 89-106). Cambridge University Press.

Vermunt, J. K., & Magidson, J. (2016). *Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax*. Statistical Innovations Inc.

Vij, A., Carrel, A., & Walker, J. L. (2013). Incorporating the influence of latent modal preferences on travel mode choice behavior. *Transportation Research Part A: Policy and Practice*, *54*, 164-178. https://doi.org/10.1016/j.tra.2013.07.008

Wanner, M., Götschi, T., Martin-Diener, E., Kahlmeier, S., & Martin, B. W. (2012). Active transport, physical activity, and body weight in adults: a systematic review. *American journal of preventive medicine, 42*(5), 493-502. https://doi.org/10.1016/j.amepre.2012.01.030

Warburton, D. E. R. (2006). Health benefits of physical activity: The evidence. *Canadian Medical Association Journal*, *174*(6), 801-809. https://doi.org/10.1503/cmaj.051351

Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020). Latent Class Analysis: A Guide to Best Practice. *Journal of Black Psychology*, *46*(4), 287-311. https://doi.org/10.1177/0095798420930932

WHO. (2022). *Walking and cycling: latest evidence to support policymaking and practice*. WHO Regional Office for Europe. ISBN: 978-92-890-5788-2.

Wikipedia. (2023, July 2). Lijst van voormalige Nederlandse gemeenten. (2023). In*Wikipedia*. https://nl.wikipedia.org/w/index.php?title=Lijst_van_voormalige_Nederlandse_gemeenten&oldid=64616364#cite_ref-MeerProv_2-62

Witte, J.-J, Alonso-González, M., & Rongen, T. (2021). *Verkenning van het concept mobiliteitshub*. Kennisinstituut voor Mobiliteitsbeleid. https://www.kimnet.nl/binaries/kimnet/documenten/rapporten/2021/05/31/verkenning-van-het-concept-mobiliteitshub/Rapport+Verkenning+Mobiliteitshub_pdfA.pdf

WMO. (2023). *State of the Global Climate 2022*. (WMO-No. 1316).

Zhang, W., Ji, C., Yu, H., Zhao, Y., & Chai, Y. (2021). Interpersonal and Intrapersonal Variabilities in Daily Activity-Travel Patterns: A Networked Spatiotemporal Analysis. *ISPRS International Journal of Geo-Information*, *10*(3), 148. https://doi.org/10.3390/ijgi10030148

Zijlstra, T., & Bingyuan, H. (2023). *Online beïnvloeden van het reisgedrag*. Kennisinstituut voor Mobiliteitsbeleid. https://www.kimnet.nl/binaries/kimnet/documenten/publicaties/2023/01/31/online-beinvloeden-van-het-reisgedrag/KiM+achtergrondrapport+Online+beinvloeden+van+het+reisgedrag_DEF.pdf

Appendix

A. Appendix: Supplemental data processing

This appendix describes in more depth the dataset explanation (Appendix A.1), dataset filtering (Appendix A.2) and dataset aggregation (Appendix A.3) for section 4.6 of Chapter 4.

A.1 Dataset explanation

Before filtering and aggregating the data, it is essential to know that the OViN data comprises information for every stage, part of a trip, in every row (one record). So, one individual can have multiple records when having more than one trip or in one trip multiple stages. Also, the matter of travel is recorded. The previous Figure 8 (Chapter 4) is extended to explain how the data is stored with certain variables, which are stored in a hierarchical structure. This means that trips cover stages, and the variable ‘has been away’ covers the stages (and trips) variables.

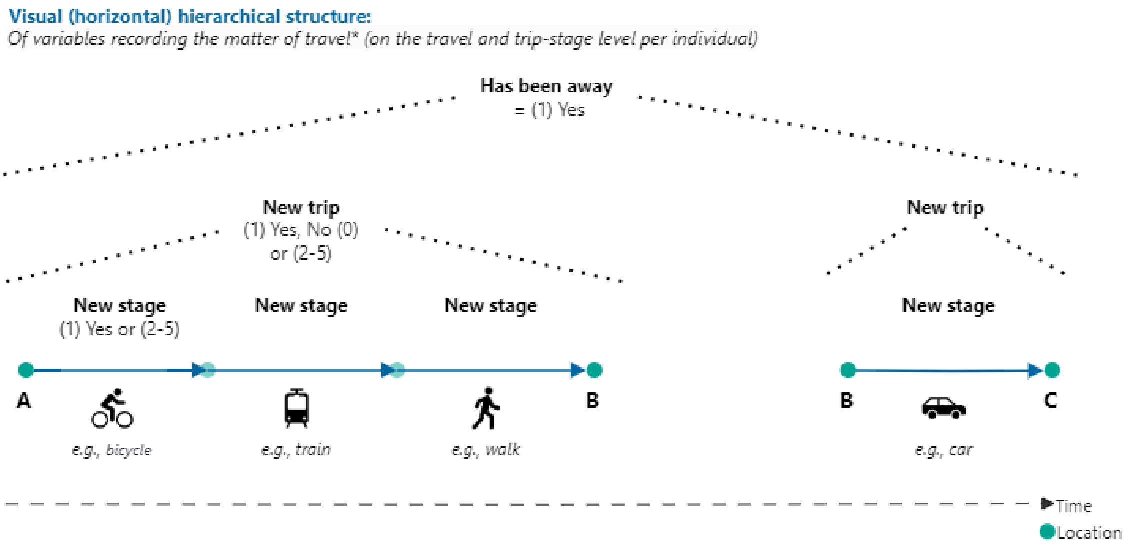
An example is given of someone who has been away in a visual horizontal structure and a corresponding tabulated vertical structure (Figure 21). This person has a personal ID number (OViN variable, in Dutch: OPID). The record number is just the general numbering of records, not part of the dataset itself. Variables and their values recording the matter of being away, a trip or stage (recorded per stage for every individual in a record) are shown; see the variables in the visualisation and table and the matters in the blue box. Other variables not included in this figure are also present in the dataset, like personal characteristics shown for every record.

Per individual, it is recorded in a variable translated to ‘has been away’ (OViN variable, in Dutch: Weggeweest) about the matter of all trips, which means that this variable stores the matter. If the matter is (2-5), all values for the has been away, trips and stages variables have correspondingly only (2), (3), (4) or (5) in the records on the corresponding variables. Trips can, however, have the same corresponding value, or a (0); this is explained later. See the explanation of the matter values in the blue box in Figure 21. The example is for someone who has been away (1), so all records have this value for the has been away variable (see the table). When this is the case, not necessarily all trips and stages variables have the yes (1) value for all records, but this is possible. So, when the has been away variable is a (1), it could be that some trips have value (1), and some others have value combinations (0-5). These trips and stages variables are now explained.

So, next to the variable ‘has been away’, the already mentioned ‘trip variable’, officially called ‘New trip’ (OViN variable, in Dutch: Verpl), shows information about a record, which is a (potential) trip. If 'New trip' is of matter (2-5), all stages within that trip have correspondingly only a (2), (3), (4), or (5) for the records on the stages variable, which is not seen as daily travel (e.g., 5 = holiday travel). Trips can, however, have the same corresponding value, or a (0). This is explained later, like is the case for the ‘has been away’ variable. Having the same values for every record on the stage variable, like the value on the trip variable, does not hold for values (1) and (0) on the trip variable.

The previous can be explained by another variable, already mentioned as the ‘stage variable’, officially called ‘New stage’ (OViN variable, in Dutch: Rit), which records the matter of the stage, where stages are part of a trip. If 'New trip' is a Yes (1) or No (0), stages can be of matter combinations (1-5) or all from matter (1). The ‘New trip’ with value (0) means that a record is at least the second stage of a trip. Several stages can thus comprise different matters, possibly from 1-5. Nevertheless, it is also possible, with a Yes (1) value on the trip variable, that all trips and stages are all also of matter (1), meaning that it is a ‘normal’ matter (not matters 2-5) and all trips comprise only one stage (otherwise value 0 is used for the trip variable).

Combining all the previous in an example, for instance, records number 1-3 in the table (Figure 21), shows that all records have a (1) on has been away; the first stage of the first trip has a (1) on the new trip variable, and the other stages of the first trip have a (0) on this variable; the stages of the second trip (1), have all (one stage) value (1) on the new trip variable. Moreover, the stages for all the trips have different matter combinations on the new stage variable.



Tabulated (vertical) hierarchical structure:
Numerical example of visual (horizontal) structure

Record Number	Personal ID	Has been away	New trip	New stage
		Matter of travel, a trip or a stage:*		
1	0	(1)	(1)	(3)
2	0	(1)	(0)	(1)
3	0	(1)	(0)	(1)
4	0	(1)	(1)	(1)

**Matter of travel, a trip or a stage:*

(0) No (not for stages = missing value)
(1) Yes
(2) Virtue of one's profession
(3) Foreign
(4) Virtue of the profession lorry drive
(5) Holiday

Hierarchical structure explained:
Of variables recording the matter of travel (per stage for every individual)*
If 'Has been away' is of matter (2-5), all records for this variable, trips**, and stages have correspondingly only a (2), (3), (4), or (5).
*** Trips can have the same value, but can also have a (0), see note.*
If 'Has been away' is of matter (1), all records for this variable have a (1). All trips can have a (1)/(0) or different matter combinations (0-5).

If 'New trip' is of matter (2-5), all records for the stages within that trip have correspondingly only a (2), (3), (4), or (5).
If 'New trip' is of matter Yes (1) or No (0), all records for stages within that trip can have a (1) or have different matter combinations (1-5).
Note: 'New trip' with value No (0) means that a record is at least the second stage of a trip.

Figure 21. The hierarchical structure of variables recording the matter of travel (per stage for every individual).

A.2 Dataset filtering

After the data is operationalised (section 4.5) and the dataset structure is explained here (Appendix A.1), some individuals can be filtered out based on their values on certain variables which do not match with the chosen sample (section 4.3) or travel behaviour definition (section 4.4). Examples of the dataset structure are also provided based on the earlier explanation (Figure 21). The excluded individuals are explained based on several variables about the sample (children) and general aspects of travel behaviour (has not been away, only non-daily travel (including foreign travel).

For other individuals who need to be filtered out, existing or to-be-made (conditional) variables are used based on values at the stage level. This is because individuals can have (some) stages which do not correspond to our travel behavioural definition. When aggregating the individual's values on records of (conditional made) variables to obtain one record per person, they can be excluded once easily. So, the obtained values of (new conditional) variables after aggregation can be used for filtering or creating new measurements. The conditional variables are made to match the definition of general aspects of travel behaviour: partly non-daily travel (including foreign travel), and the travel behavioural mode set aspect (another mode set). Aggregation is further explained in Appendix A.3. Several measurements (travel behaviour and multimodality measurements) are created after aggregation based on the used and created variables, explained in section 4.7.

The excluded individuals are explained based on several variables about the sample (children) and general aspects of travel behaviour (has not been away, only non-daily travel (including foreign travel). Conditional variables are made to match the definition of general aspects of travel behaviour (partly non-daily travel (including foreign travel); partly non-daily travel (foreign travel)) and travel behavioural mode set (another mode set). Foreign travel has several options in being present in the data, so it is explained in several parts.

Children, has not been away, only non-daily travel (including foreign travel)

Regarding the defined sample (adults) and the travel behaviour itself (having travelled with daily travel), some individuals (all of their records) are excluded and filtered out (children, has not been away and non-daily travel). First, records for individuals below age 18 have been excluded (OViN variable, in Dutch: Leeftijd); see record numbers 6-8 as an example which all have a value below 18 for this individual (ID 2) in Figure 22. Second, individuals who have not been away in the diary period (0) (ID 1) or made only trips which are not seen as daily travel, i.e. non-daily travel (2-5) (ID 2), are excluded. The latter individuals include only having trips made by virtue of one's profession (e.g., as a courier), for foreign trips, by virtue of the profession of lorry driver or for holiday travel. These non-travellers and non-daily travellers are excluded by only keeping the records with a (1) on the variable 'has been away', meaning that they have the value 'yes' and thus did travel. For instance, record 5 is excluded as it has a (0) on the variable 'has been away', so it has not been away. Another example is that records 6/7 and 8 are excluded as all records have a (2) on the variable 'has been away', although this person is also excluded based on age.

Because the mentioned variables ('age' and 'has been away') have the same value for every individual record, the exclusion criteria result in keeping only individuals with all their corresponding records who are 18 or older, have been away and not did only non-daily travel.

		Matter of travel, a trip or a stage:*			
Record Number	Personal ID	Has been away	New trip	New stage	Age
5	1	(0)	(0)	.	63
6	2	(2)	(2)	(2)	16
7	2	(2)	(0)	(2)	16
8	2	(2)	(2)	(2)	16

*See Figure 21.

Figure 22. Example in the tabulated form of individuals excluded (children, has not been way and non-daily travel).

Partly non-daily travel (including foreign travel)

Due to the hierarchical set-up of has been away, trips and stages, it should be accounted for that if someone has been away (1), several trips or stages of different kinds might occur, which are not seen as daily travel. This is shown in the example in Figure 23, where the same individual (ID 3) makes trips and stages of kind 1 (daily travel) and kind 5. Moreover, if a trip is recorded as matter (1), it might be that this trip comprises several stages, where some stages are of kind (1), and another stage is of another matter when one trip had combined kinds for stages. This is visualised in Figure 24, which also shows that part of this travel is foreign travel. This person (ID 4) travelled in a foreign country to the airport (for instance) with a distance of 50 km (500 hm). The next stage consists of travelling by aeroplane (for instance) for 1000 km (10000 hm) abroad and 50 km (500 hm) in The Netherlands before arriving at the airport. The next stage consisted of travelling 25 km (250 hm) to the end destination in The Netherlands. Distance tracking is further explained in the upcoming part.

A new variable, 'Stage delete', has been created as a conditional variable. Records got a one when they do have stages which are of kinds 2-5 (after filtering on has been away), to account for all the possibilities of having (partly) trips or stages of these matters (including foreign travel).

		Matter of travel, a trip or a stage:*		
Record Number	Personal ID	Has been away	New trip	New stage
9	3	(1)	(1)	(1)
10	3	(1)	(0)	(1)
11	3	(1)	(5)	(5)
12	3	(1)	(0)	(5)

*See Figure 21.

Figure 23. Example in the tabulated form of individuals excluded (partly non-daily travel).

		Matter of travel, a trip or a stage:*				
Record Number	Personal ID	Has been away	New trip	New stage	Distance** NL	Distance** abroad
13	4	(1)	(1)	(3)	0	500
14	4	(1)	(0)	(1)	500	10000
15	4	(1)	(0)	(1)	250	0

*See Figure 21. **Measured in hm.

Figure 24. Example in tabulated form of individuals excluded (partly non-daily travel, including foreign travel).

Partly non-daily travel (foreign travel)

Someone who is entirely travelling abroad is already filtered out, as explained, as well as someone partly travelling abroad. Individuals who have stages with a (1) are left in the data based on the previous filtering and use of conditional variables. However, foreign trips are possible for all stage coding (1-5), also when the trip is not yet filtered out based on previous conditions and the stage is not marked as a complete foreign stage (3). In order to exclude the individuals who are partly abroad travelling, an example is given in Figure 25, based on variables which are tracking the distance travelled per stage in The Netherlands, 'distance NL' (OViN variable, in Dutch: AfstR) or abroad 'distance abroad' (OViN variable, in Dutch: AfstRBL). The imaginary person (ID 5) travelled abroad for 15 km (150 hm) in The Netherlands, and 25 km (250 hm) abroad, for the first trip doing groceries abroad. Then 25 km (250 hm) abroad are travelled when going back, and 10 km (100 hm) are travelled in The Netherlands for a visit, for instance. Then the third stage is 5 km (50 hm) in The Netherlands when going home.

The existing variable 'distance abroad' (see official name before) tracks the amount of foreign travel in a stage. When having a value bigger than zero hm, this variable can be used as a conditional variable to filter out individuals who still have foreign travel in their travel behaviour.

		Matter of travel, a trip or a stage:*				
Record Number	Personal ID	Has been away	New trip	New stage	Distance** NL	Distance** abroad
16	5	(1)	(1)	(1)	150	250
17	5	(1)	(0)	(1)	100	250
18	5	(1)	(0)	(1)	50	0

*See Figure 21. **Measured in hm.

Figure 25.Example of individuals excluded (partly foreign travel).

Other mode set

When examining travel behaviour, only individuals are examined if they only had a travel pattern exclusively with the set of specified modes. In the operationalisation, these are marked in categories of the variable tracking the stage mode (OViN variable, in Dutch: Rvm, transformed in Table 9). A new variable, 'mode delete', is created, where every record obtains a one if the used mode for that stage is out of consideration, which can be used as a conditional variable.

A.3 Dataset aggregation

As it is explained which data should be filtered before and after aggregation before, based on (newly created) variables, the aggregation itself is further explained. But first, the OViN data for all the years are combined into one dataset. For the aggregation itself, the 'break' variable for aggregating on the induvial level is 'OPID' (OViN variable, personal ID) to ensure that all records (containing one stage) per individual are used to aggregate. This implies that after aggregation, every individual contains one record (row) instead of multiple (one record per stage). Before aggregation, the data is filtered for children, individuals who have not been away, or only did non-daily travel.

Regarding the previously explained (existing) conditional variables in the data filtering section (Appendix A.2), to obtain the correct data according to the definition of travel behaviour, individuals are included/excluded based on these (created) variables and aggregation of them. The variable 'Stage delete' is summated over every individual record during aggregation. After aggregation, individuals having a one or higher have stages which are not marked as daily travel. So, individuals are included when having a zero on this variable. The existing variable 'distance abroad' is also summated over every individual record during aggregation. After aggregation, individuals having not travelled abroad at no stage have a zero on this variable and are included. The variable 'mode delete' is also summated for every individual based on their records. When having a one or higher on this variable, these individuals used for one or multiple stages modes outside the defined mode set. Only individuals having a zero on this variable are included.

For the active covariates (Table 8), the first record is kept, as this variable is the same for every record (for every individual). This is also the case for the inactive covariate residential municipality. For the travel behavioural variables and multimodality measurements, section 4.7 describes how these are measured and created (with the help of variables) during aggregation.

B. Appendix: Descriptive statistics

This appendix shows the representative analysis (Table 19), the distributions of the sample for age, education level, household income and urban density (Figure 26, Figure 27, Figure 28, Figure 29), and the correlation of the active covariates (Figure 30) is shown for section 4.8 of Chapter 4.

Representative analysis

Table 19. Representative analysis sample compared to the target population.

Variable		Sample	CBS*	
Active covariates				
Socio-demographic variables				
Gender (%)	Male	47	49	a
	Female	53	51	
Age (%)	18 - 32	20	23	b
	33 - 44	20	20	
	45 - 54	20	19	
	55 - 64	19	16	
	65+	21	21	
Ethnicity (%)	Native	85	79	c
	Immigrant	15	21	
	Missing	~0	0	
Education level (%)	No education	0.7	0.0	d
	Primary education	4.3	10.0	
	Pre-vocational education	21.5	23.0	
	Vocational & higher secondary education	38.0	38.0	
	Higher vocational & university education	33.8	28.0	
	Missing	1.7	1.0	
Household composition (%)	One-person household	18	37	e
	Couple household	36	29	
	Other household	47	34	
Occupation (%)	Employed	57	66	f
	Student	6	-	
	Other occupation	34	-	
	Missing	3	-	
Household income (%) (standardised)	< €10,000	3	5	g
	€10,000 - €20,000	26	30	
	€20,000 - €30,000	38	34	
	€30,000 - €40,000	20	19	
	€40,000 - €50,000	7	7	
	≥ €50,000	5	5	
	Missing	~0	0	
Mobility resource variables				
Licensure (%)	No	12	20	h
	Yes	88	80	
	Missing	~0	-	
Ownership	No	11	29	i

household car (%)	Yes	89	71	
Ownership company car (%)	No	94	89	j
	Yes	6	11	
Ownership household bicycle (%)	No	5	-	
	Yes	95	-	
Built-environment variables				
Urban density (%)	High (≥ 1500 addresses/km^2)	45	45	k
	Medium (1000-1500 addresses/km^2)	18	19	
	Low (< 1000 addresses/km^2)	37	36	

*Data or calculations (merging categories) are not entirely similar to our data: our sample is 18+ (not always the case for CBS data), our data is 2010-2017 (not always the whole period for CBS data), and some categories or definitions are not fully comparable.
Blue = categories/definitions of CBS are likely quite similar to our operationalisation.

Data tables (online customised) based on Statistics Netherlands (CBS):
a: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/7461BEV/table?dl=93696
b: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/7461BEV/table?dl=9369B
c: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70072ned/table?dl=93698
d: https://opendata.cbs.nl/#/CBS/nl/dataset/82275NED/table?dl=945CC
e: https://opendata.cbs.nl/#/CBS/nl/dataset/37975/table?dl=9369A
f: https://opendata.cbs.nl/#/CBS/nl/dataset/80590ned/table?dl=9369D
g: https://opendata.cbs.nl/#/CBS/nl/dataset/83932NED/table?dl=936A0
h: https://opendata.cbs.nl/#/CBS/nl/dataset/83488NED/table?dl=936A1
i: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81845NED/table?dl=936A2
j: https://opendata.cbs.nl/#/CBS/nl/dataset/71405ned/table?dl=936A3
k: https://opendata.cbs.nl/#/CBS/nl/dataset/70072ned/table?dl=9369F
-: information not available

Distributions

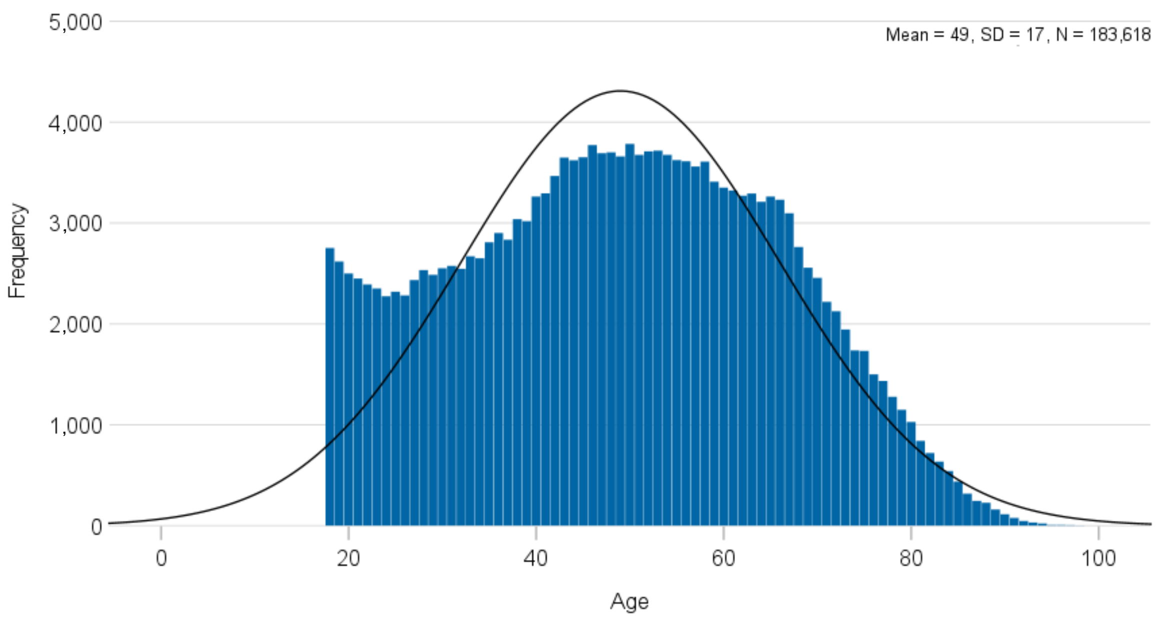
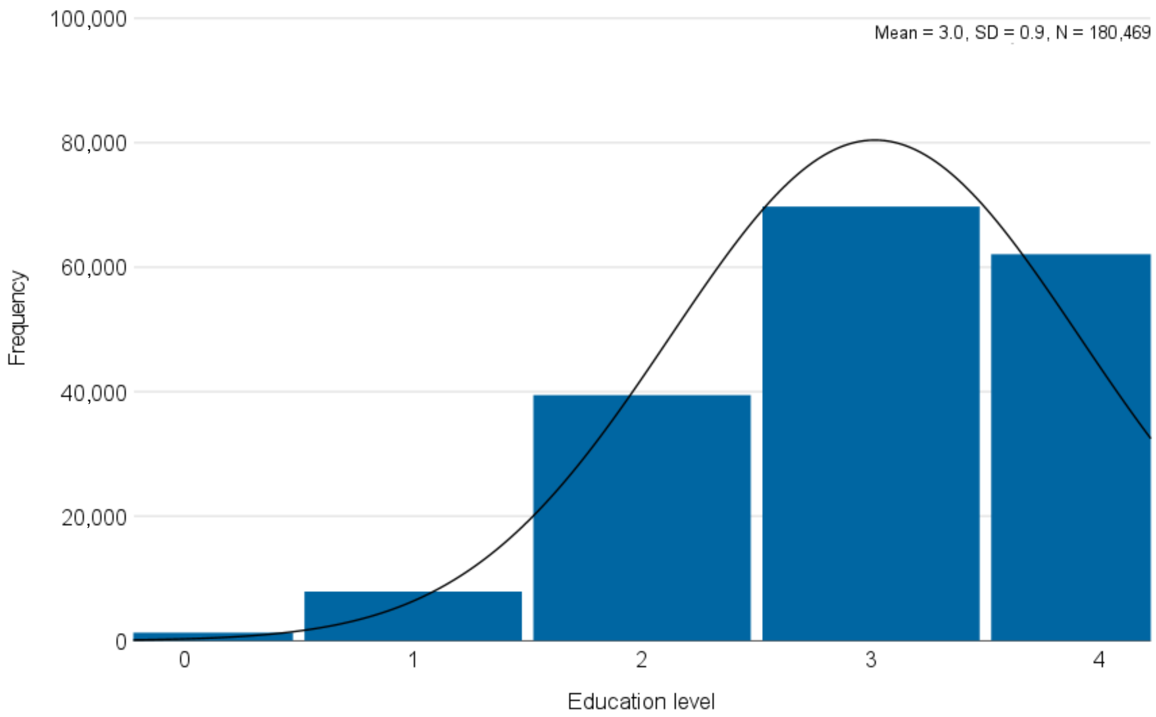
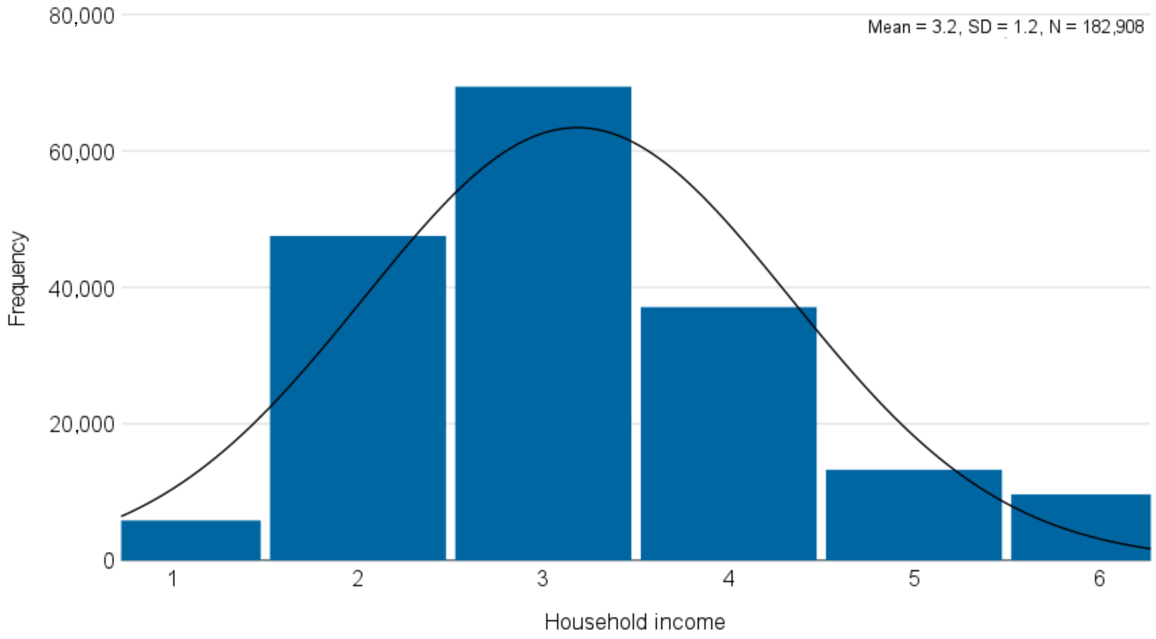


Figure 26. Frequency distribution age (and normal distribution line).



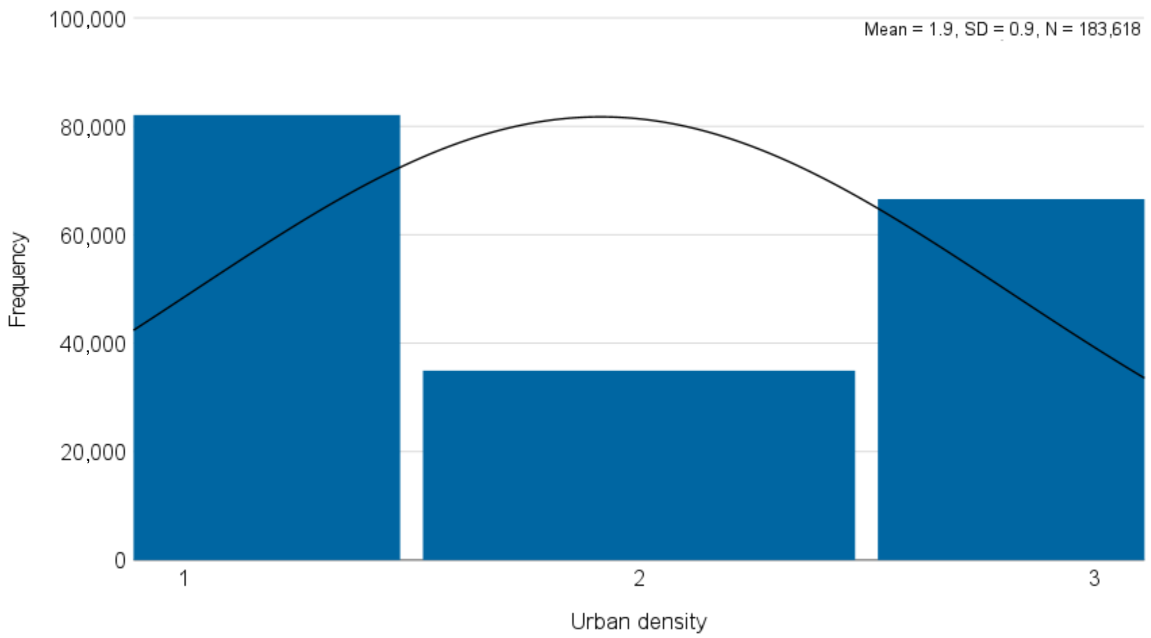
No education (0), Primary education (1), Pre-vocational education (2), Vocational & higher secondary education (3), Higher vocational & university education (4)

Figure 27. Frequency distribution education level (and normal distribution line).



< €10,000 (1), €10,000 - €20,000 (2), €20,000 - €30,000 (3), €30,000 - €40,000 (4), €40,000 - €50,000 (5), ≥ €50,000 (6)

Figure 28. Frequency distribution household income (level) (and normal distribution line).



High (≥ 1500 addresses/km^2) (1), Medium (1000-1500 addresses/km^2) (2), Low (< 1000 addresses/km^2) (3)

Figure 29. Frequency distribution urban density (class) (and normal distribution line).

Correlations

Correlation between variables	Socio-demographics													Mobility resources								Built-environment
	Gender _male	Gender _female	Age	Ethnicity _native	Ethnicity _immigrant	Education level	Household composition _one person	Household composition _couple	Household composition _other	Occupation _employed	Occupation _student	Occupation _other	Household income	Licensure _no	Licensure _yes	Household car_no	Household car_yes	Company car_no	Company car_yes	Household bicycle_no	Household bicycle_yes	Urban density*
Socio-demographics																						
Gender_male																						
Gender_female																						
Age	0.01	-0.01																				
Ethnicity_native	0.00	0.00	0.09																			
Ethnicity_immigrant	0.00	0.00	-0.09																			
Education level	0.06	-0.06	-0.21	0.02	-0.02																	
Household composition_one person	-0.02	0.02	0.12	-0.01	0.01	-0.05																
Household composition_couple	0.03	-0.03	0.40	0.08	-0.08	-0.07																
Household composition_other	-0.01	0.01	-0.48	-0.07	0.07	0.10																
Occupation_employed	0.09	-0.09	-0.40	0.01	-0.01	0.31	-0.12	-0.22	0.30													
Occupation_student	0.00	0.00	-0.42	-0.05	0.05	-0.03	0.00	-0.16	0.15													
Occupation_other	-0.10	0.10	0.63	0.01	-0.01	-0.30	0.12	0.31	-0.38													
Household income	0.04	-0.04	0.05	0.08	-0.08	0.30	-0.17	0.14	-0.01	0.21	-0.11	-0.16										
Mobility resources																						
Licensure_no	-0.12	0.12	0.00	-0.12	0.12	-0.23	0.12	-0.06	-0.04	-0.23	0.19	0.14	-0.19									
Licensure_yes	0.12	-0.12	0.00	0.12	-0.12	0.23	-0.12	0.06	0.04	0.23	-0.19	-0.14	0.19									
Household car_no	-0.05	0.05	-0.01	-0.11	0.11	-0.10	0.39	-0.14	-0.16	-0.16	0.13	0.10	-0.25	0.45	-0.45							
Household car_yes	0.05	-0.05	0.01	0.11	-0.11	0.10	-0.39	0.14	0.16	0.16	-0.13	-0.10	0.25	-0.45	0.45							
Company car_no	-0.16	0.16	0.06	-0.03	0.03	-0.12	0.04	0.04	-0.07	-0.18	0.06	0.16	-0.15	0.09	-0.09	0.08	-0.08					
Company car_yes	0.16	-0.16	-0.06	0.03	-0.03	0.12	-0.04	-0.04	0.07	0.18	-0.06	-0.16	0.15	-0.09	0.09	-0.08	0.08					
Household bicycle_no	-0.02	0.02	0.16	-0.08	0.08	-0.14	0.20	0.00	-0.15	-0.14	-0.03	0.16	-0.10	0.12	-0.12	0.14	-0.14	0.03	-0.03			
Household bicycle_yes	0.02	-0.02	-0.16	0.08	-0.08	0.14	-0.20	0.00	0.15	0.14	0.03	-0.16	0.10	-0.12	0.12	-0.14	0.14	-0.03	0.03			
Built-environment																						
Urban density*	0.00	0.00	0.07	0.18	-0.18	-0.10	-0.11	0.05	0.04	-0.02	-0.05	0.05	0.00	-0.10	0.10	-0.19	0.19	0.01	-0.01	-0.06	0.06	
Time																						
Year	0.00	0.00	0.05	-0.03	0.03	0.04	0.03	0.00	-0.03	-0.02	0.00	0.02	0.13	0.00	0.00	0.03	-0.03	0.00	0.00	0.03	-0.03	-0.10
Year_2010	0.00	0.00	-0.03	0.02	-0.02	-0.02	-0.02	0.00	0.01	0.01	0.00	-0.02	-0.04	0.00	0.00	-0.02	0.02	0.00	0.00	-0.03	0.03	0.04
Year_2011	0.00	0.00	-0.03	0.01	-0.01	-0.02	-0.01	-0.01	0.02	0.02	0.00	-0.02	-0.04	0.01	-0.01	-0.01	0.01	0.00	0.00	-0.02	0.02	0.03
Year_2012	0.00	0.00	-0.01	0.01	-0.01	-0.01	-0.01	0.00	0.00	0.01	-0.01	0.00	-0.03	0.00	0.00	-0.01	0.01	0.00	0.00	-0.02	0.02	0.02
Year_2013	0.00	0.00	0.00	0.01	-0.01	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	-0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.03	-0.03	0.03
Year_2014	0.00	0.00	0.02	0.01	-0.01	0.00	0.00	0.00	0.00	-0.01	0.00	0.01	-0.02	-0.01	0.01	-0.01	0.01	0.00	0.00	0.05	-0.05	0.03
Year_2015	0.00	0.00	0.01	-0.02	0.02	0.01	0.01	0.00	-0.01	-0.01	0.00	0.01	-0.01	0.01	-0.01	0.02	-0.02	-0.01	0.01	0.01	-0.01	-0.06
Year_2016	0.00	0.00	0.01	-0.02	0.02	0.02	0.01	0.00	-0.01	0.00	0.00	0.00	0.08	0.00	0.00	0.01	-0.01	0.00	0.00	0.00	0.00	-0.06
Year_2017	0.00	0.00	0.03	-0.02	0.02	0.02	0.01	0.01	-0.02	-0.01	0.00	0.01	0.11	0.00	0.00	0.01	-0.01	0.00	0.00	-0.01	0.01	-0.06

Notes: Bivariate correlations (significance is not shown), missing values are pairwise excluded, correlations between (categories of) the same variable are not shown, the discrete colour scale is used per decimal category from light blue (0.00) to mid blue (≤ 0.49 or ≥ -0.49), and the colour darker blue is used (≥ 0.50 or ≤ -0.50).
*A higher value means a lower urban density category.

Figure 30. Correlation between active covariates (zoom in for readability).

C. Appendix: Model outcome

This appendix shows the latent class model outcomes, belonging to the results of Chapter 6: latent class model parameters (Table 20) and corresponding obtained latent class profiles (Table 21) and latent class membership (Table 22).

Latent class model parameters

Table 20. Parameters 5-class solution: prediction for indicators and latent class membership (effect coding) from the latent class model with indicators and covariates.

	Class[a]		(1) C	(2) B	(3) CW	(4) CB MM	(5) PT+ MM	Wald
	Inter-cept	Wald						
Model for indicators[b] (measurement model)								
# Walk stages	-0.63	5386*	-2.53	-0.19	1.40	-0.32	1.64	54235*
# Bicycle stages	-3.23	51*	-8.71	4.22	-2.73	3.95	3.27	12041*
# PT stages	-4.55	4049*	-1.55	-2.15	-0.55	-1.08	5.34	8445*
# Car stages	-1.95	19*	3.05	-9.48	2.31	2.80	1.32	29940*
Model for covariates (structural model)								
Intercept			0.78	-0.30	1.05	-2.00	0.48	1749*
Gender								792*
Male			0.12	0.01	-0.06	-0.09	0.02	
Female			-0.12	-0.01	0.06	0.09	-0.02	
Age			-0.004	0.008	0.003	0.005	-0.012	705*
Ethnicity								815*
Native			-0.07	0.10	-0.05	0.21	-0.18	
Immigrant			0.07	-0.10	0.05	-0.21	0.18	
Education level			-0.15	-0.05	-0.07	0.08	0.19	919*
Household composition								180*
One-person household			0.05	-0.05	0.07	-0.06	-0.01	
Couple household			-0.03	0.07	-0.01	-0.04	0.02	
Other household			-0.01	-0.01	-0.07	0.10	-0.01	
Occupation (%)								3892*
Employed			0.32	-0.16	0.07	0.05	-0.27	
Student			-0.45	0.25	-0.58	-0.09	0.87	
Other occupation			0.13	-0.09	0.51	0.04	-0.59	
Household income (standardised)			0.01	-0.06	-0.07	0.01	0.10	341*
Licensure (%)								2955*
No			-0.46	0.19	0.11	-0.27	0.42	
Yes			0.46	-0.19	-0.11	0.27	-0.42	
Ownership household car								5225*
No			-0.84	0.49	0.10	-0.34	0.60	
Yes			0.84	-0.49	-0.10	0.34	-0.60	
Ownership company car								1600*
No			-0.44	0.42	-0.33	-0.13	0.48	
Yes			0.44	-0.42	0.33	0.13	-0.48	
Ownership household bicycle								2569*
No			0.45	-0.73	0.43	-0.39	0.25	
Yes			-0.45	0.73	-0.43	0.39	-0.25	
Urban density**			0.18	0.04	0.01	0.16	-0.32	1953*

Table 20. (continued). Parameters 5-class solution: prediction for indicators and latent class membership (effect coding) from the latent class model with indicators and covariates.

Class[a]	(1) C	(2) B	(3) CW	(4) CB MM	(5) PT+ MM	Wald
Model for covariates (structural model) cont'd						
Year						290*
2010	-0.02	-0.10	0.14	0.01	-0.03	
2011	-0.01	-0.04	-0.05	0.05	0.06	
2012	0.02	-0.01	-0.04	0.01	0.01	
2013	-0.02	0.00	0.00	-0.01	0.03	
2014	-0.07	0.09	-0.09	0.05	0.02	
2015	0.02	0.01	-0.01	0.00	-0.01	
2016	0.03	0.02	0.06	-0.04	-0.06	
2017	0.06	0.03	0.00	-0.07	-0.01	

a: (1) C: Car exclusive, (2) B: Bicycle mostly, (3) CW: Car + Walk, (4) CB MM: Car + Bicycle Multimodal, (5) PT+ MM: Public transport+ Multimodal.
b: Indicators are measured over one day.
* Significant on the 95% confidence interval (p < 0.00).
** A higher value means a lower urban density category.

Note: Example interpretation:
For (intercept of) indicators parameters and covariate intercept parameters interpretation example, see Ton et al. (2020, p. 1851)
A female is less likely (par. is -0.12) to be in the (1) class, compared to the average (effect coding is used).
Moreover, a female is more associated with being in the (4) class (par. is 0.09)

Latent class profiles

Table 21. Profiles 5-class solution: within-class distribution/values from the latent class model with indicators and covariates.

	Class[a]	(1) C	(2) B	(3) CW	(4) CB MM	(5) PT+ MM	Total sample [b]
Class size (%) N = 183,618		**41**	20	17	12	10	100
Indicators							
Travel behaviour							
# Walk stages	Mean	0.0	0.4	**2.2**	0.4	**2.8**	0.8
# Bicycle stages	Mean	0.0	**2.7**	0.0	**2.0**	1.0	0.9
# PT stages	Mean	0.0	0.0	0.0	0.0	**2.2**	0.2
# Car stages	Mean	**3.0**	0.0	1.4	**2.3**	0.5	1.8
*Total stages**	*Sum of means*	*3.0*	*3.1*	*3.6*	*4.7*	*6.5*	*3.7*
Covariates							
Socio-demographic variables							
Gender (%)							
Male		**53**	43	40	42	44	47
Female		47	57	**60**	58	56	53
Age (%)							
18-32		18	19	15	16	**45**	20
33-44		**23**	16	18	**23**	16	20
45-54		**23**	19	17	**23**	15	20
55-64		19	**21**	19	20	13	17
65+		18	**25**	**31**	18	11	21
	Mean	49	**51**	**54**	49	39	49
Ethnicity (%)							
Native		86	87	84	**92**	73	85
Immigrant		14	13	16	8	**27**	15
Missing		0	0	0	0	0	~0
Education level (%)							
No education		1	1	1	0	1	0.7
Primary education		3	5	**7**	2	3	4.3
Pre-vocational education		21	**25**	**25**	19	15	21.5
Vocational & higher secondary ed.		**40**	35	35	38	**41**	38.0
Higher vocational & university ed.		34	32	30	**39**	**38**	33.8
Missing		2	2	2	2	2	1.7
Household composition (%)							
One-person household		14	21	23	12	**25**	18
Couple household		36	38	**40**	36	23	36
Other household		**50**	41	38	**52**	**52**	47
Occupation (%)							
Employed		**67**	48	45	**62**	51	57
Student		3	8	2	4	**27**	6
Other occupation		28	**41**	**50**	31	20	34
Missing		3	3	3	3	2	3

Table 21. (continued). Profiles 5-class solution: within-class distribution from the latent class model with indicators and covariates.

Class[a]	(1) C	(2) B	(3) CW	(4) CB MM	(5) PT+ MM	Total sample [b]
Socio-demographic variables (continued)						
Household income (standardised) (%)						
< €20,000	24	**35**	**35**	23	**35**	29
€20,000 - €30,000	39	37	37	**40**	34	38
€30,000 - €40,000	**22**	18	17	**23**	19	20
≥ €40,000	**14**	10	10	**14**	12	12
Missing	0	1	0	0	1	~0
Mobility resource variables						
Licensure (%)						
No	3	21	17	5	**32**	12
Yes	**97**	79	83	**95**	68	88
Missing	0	0	0	0	0	~0
Ownership household car (%)						
No	1	23	15	3	**34**	11
Yes	**99**	77	85	**97**	67	89
Ownership company car (%)						
No	90	**99**	95	95	**99**	94
Yes	**10**	1	5	5	1	6
Ownership household bicycle (%)						
No	6	1	**11**	1	7	5
Yes	94	**99**	89	**99**	93	95
Built-environment variables						
Urban density (%)						
High (≥ 1500 addresses/km^2)	38	49	47	37	**68**	45
Medium (1000-1500 addresses/km^2)	20	19	19	**21**	14	19
Low (< 1000 addresses/km^2)	**42**	31	34	**42**	18	36
Time						
Year (%)						
2010	13	12	15	14	12	13
2011	13	12	12	14	13	13
2012	14	13	13	14	13	13
2013	13	13	13	13	13	13
2014	13	14	13	14	13	13
2015	11	12	12	11	12	12
2016	11	12	12	11	12	11
2017	12	12	11	11	12	12

Table 21. (continued). Profiles 5-class solution: within-class distribution from the latent class model with indicators and covariates.

	Class[a]	(1) C	(2) B	(3) CW	(4) CB MM	(5) PT+ MM	Total sample [b]
Inactive covariates							
Multimodality measurements of travel behaviour (over one day)							
Number of modes (%)							
1		96	78	53	16	11	
2		4	22	46	67	48	
3		0	0	1	17	36	
4		0	0	0	0	5	
	Mean	1.0	1.2	1.5	2.0	**2.4**	1.4
OM_PI (%)							
0		96	78	53	16	11	
0.16-0.46		4	12	21	31	26	
0.46-1.00		0	10	25	53	63	
	Mean	0.02	0.10	0.22	**0.44**	**0.54**	0.2
Built-environment variables							
Residential municipality (%)					-		

Note: Some column values of (in)active covariates may not add up to 100% due to rounding
a: (1) C: Car exclusive, (2) B: Bicycle mostly, (3) CW: Car + Walk, (4) CB MM: Car + Bicycle Multimodal, (5) PT+ MM: Public transport+ Multimodal.
b: Data are pooled for 2010-2017.
* Calculated on own.
Bold = The highest sizes/means/shares for a (category of) variable compared to other classes.

Latent class membership

Table 22. ProbMeans 5-class solution: outside-class probability distribution from the latent class model with indicators and covariates.

Class[a]	(1) C	(2) B	(3) CW	(4) CB MM	(5) PT+ MM
Class size (%) N = 183,618	41	20	17	12	10
Indicators					
Travel behaviour (over one day)					
# Walk stages					
0	59	22	3	14	3
1	18	24	33	17	8
3-16	1	12	51	6	30
# Bicycle stages					
0	63	1	26	2	8
1-2	0	51	0	35	13
3-17	0	66	0	23	10
# PT stages					
0	45	22	18	13	2
1-14	1	0	1	0	97
# Car stages					
0	0	56	21	3	20
1	26	0	24	29	21
2	62	0	16	18	4
3	69	0	14	15	3
4-16	74	0	11	13	1
Covariates					
Socio-demographic variables					
Gender (%)					
Male	47	18	15	11	9
Female	36	21	19	13	10
Age (%)					
18-32	38	19	12	10	22
33-44	47	16	15	14	8
45-54	46	18	15	14	7
55-64	41	22	17	13	7
65+	35	24	26	11	5
Ethnicity (%)					
Native	42	20	17	13	8
Immigrant	39	18	19	7	18
Missing	30	26	12	2	30

Table 22. (continued). ProbMeans 5-class solution: outside-class probability distribution from the latent class model with indicators and covariates.

Class[a]	(1) C	(2) B	(3) CW	(4) CB MM	(5) PT+ MM
Socio-demographic variables (continued)					
Education level (%)					
No education	35	22	27	4	12
Primary education	34	25	28	7	7
Pre-vocational education	40	23	20	11	7
Vocational & higher secondary education	44	18	16	12	10
Higher vocational & university education	41	18	15	14	11
Missing	39	22	20	11	9
Household composition (%)					
One-person household	32	24	22	8	14
Couple household	42	21	19	12	6
Other household	45	17	14	14	11
Occupation (%)					
Employed	48	17	13	13	9
Student	17	26	6	8	43
Other occupation	34	24	25	11	6
Missing	39	23	16	13	9
Household income (standardised) (%)					
< €20,000	35	24	21	9	12
€20,000 - €30,000	43	19	16	13	9
€30,000 - €40,000	45	18	14	14	9
≥ €40,000	47	15	14	14	9
Missing	26	28	17	7	21
Mobility resource variables					
Licensure (%)					
No	10	34	24	5	26
Yes	46	18	16	13	8
Missing	15	32	44	1	7
Ownership household car (%)					
No	5	40	23	3	29
Yes	46	17	16	13	7
Ownership company car (%)					
No	40	21	17	12	10
Yes	70	4	14	11	2
Ownership household bicycle (%)					
No	44	5	35	3	12
Yes	41	21	16	13	10
Built-environment variables					
Urban density (%)					
High (≥ 1500 addresses/km^2)	35	22	18	10	15
Medium (1000-1500 addresses/km^2)	43	20	17	14	7
Low (< 1000 addresses/km^2)	48	17	16	14	5

Table 22. (continued). ProbMeans 5-class solution: outside-class probability distribution from the latent class model with indicators and covariates.

Class[a]	(1) C	(2) B	(3) CW	(4) CB MM	(5) PT+ MM
Time					
Year (%)					
2010	41	18	19	13	9
2011	42	19	16	13	10
2012	42	20	16	12	9
2013	41	19	17	12	10
2014	40	21	16	13	9
2015	41	20	17	12	10
2016	41	20	18	11	10
2017	42	20	17	11	10
Inactive covariates					
Multimodality measurements of travel behaviour (over one day)					
Number of modes (%)					
1	59	23	13	3	2
2	6	16	29	31	18
3	0	0	2	37	61
4	0	0	0	2	98
OM_PI (%)					
0	59	23	13	3	2
0.16-0.46	6	17	26	28	18
0.46-1.00	1	10	23	34	32
Built-environment variables					
Residential municipality (%)		See section 5.4 in Chapter 5.			

Note: Some row values of (in)active covariates may not add up to 100% due to rounding.
a: (1) C: Car exclusive, (2) B: Bicycle mostly, (3) CW: Car + Walk, (4) CB MM: Car + Bicycle Multimodal, (5) PT+ MM: Public transport+ Multimodal.

COLOPHON

Determinants and development of multimodal travel patterns: Identifying travel user groups in The Netherlands using Latent Class Cluster Analysis

MSc Thesis by F. Verheij

2023

Master Complex Systems Engineering and Management
Faculty of Technology, Policy and Management
Delft University of Technology