Travel behaviour in suburban areas: A case study in the province of Zuid-Holland using latent class cluster analysis.

Max Klootwijk
3-12-2024

Delft University of Technology
MSc. Transport, Infrastructure and Logistics

Travel behaviour in suburban areas: A case study in the province of Zuid-Holland using latent class cluster analysis.

by

M.R. Klootwijk | 2-12-2024

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday December 12-12, 2024 at 14:45.

Student number:	4712781	
Project duration:	February, 2024 – December, 2024	
Thesis committee:	Dr. J.A. Annema,	TU Delft, Chair
	Dr. C. Maat,	TU Delft, supervisor
	Prof. Dr. G.P. van Wee,	TU Delft, (3rd) supervisor
	Dr. Ir. L.B. Zomer,	Provincie Zuid-Holland, supervisor

Preface

In de afgelopen 9 maanden heb ik hard gewerkt aan deze thesis. Mijn onderzoek kende aanvankelijk een compleet andere opzet dan het onderzoek wat nu voor u ligt. Het initiële plan was namelijk om een onderzoek uit te voeren dat gebaseerd zou zijn op enquêtes. Dit veranderde echter naar een volledig datagedreven onderzoek, wat heeft geleid tot interessante inzichten en resultaten.

Dit onderzoek heb ik individueel uitgevoerd, maar zeker niet zonder hulp. Daarom wil ik eerst mijn begeleider vanuit de provincie Zuid-Holland, Lara Zomer enorm bedanken. De vele gesprekken, feedback en discussies over de uitvoering van dit onderzoek hebben het onderzoek zeker een meerwaarde gegeven.

Gedurende mijn afstudeerstage vond ik het bovendien erg interessant om mee te lopen in het provinciehuis, een plek waar beleid wordt gemaakt en geschreven. Ik heb daar een erg leuke tijd gehad en wil daarom ook alle collega's en mede-afstudeerders van de provincie, met wie ik over mijn thesis heb kunnen sparren, enorm bedanken voor hun interesse en waardevolle input.

Vanuit de TU Delft wil ik mijn begeleiders, Kees Maat en Jan Anne Annema bedanken voor hun goede begeleiding en input tijdens onze besprekingen. Ik denk dat we met dit afstudeeronderzoek een goede weg zijn ingeslagen. Jullie feedback heeft hier zeker aan bijgedragen. Verder wil ik ook het TU Delft Writing center bedanken voor de tips bij het formuleren en structureren van de thesis.

Tot slot wil ik ook mijn ouders, zusje en vrienden bedanken voor hun steun en betrokkenheid tijdens mijn afstuderen. Zonder hen zou de eindfase van mijn studietijd aan de TU Delft zeker moeizamer zijn geweest.

Ik wens jullie veel plezier met het lezen van mijn thesis! …

M.R. Klootwijk
Delft, December 2024

Samenvatting

De provincie Zuid-Holland heeft de afgelopen 20 jaar een grote stijging gezien in zowel autobezit als -gebruik. Dit brengt diverse uitdagingen met zich mee, zoals congestie, een groot ruimtebeslag voor de auto en uitdagingen op het gebied van toekomstige woningbouw. Bovendien is de verwachting dat het autobezit in de toekomst verder zal toenemen.

Deze stijging in autobezit is vooral zichtbaar in gebieden buiten de grote steden, waardoor deze gebieden specifieke aandacht verdienen. Deze gebieden kunnen ingedeeld worden in diverse subcategorieën, zoals landelijke gebieden, maar ook gebieden tussen de stad en het landelijk gebied in. Deze laatste categorie, ook wel aangeduid als suburbaan gebied, behoeft meer verdieping vanwege de stijging in autobezit afgelopen 20 jaar. Dit is opmerkelijk, aangezien suburbane gebieden wel nabijheid hebben tot de stad, waar volop voorzieningen beschikbaar zijn. Daarnaast heeft suburbaan gebied ook alternatieven voor autogebruik, zoals OV en fiets.

Het doel van deze studie is om suburbane gebieden nader te onderzoeken. De hoofdvraag is welk type reisgedrag in suburbane gebieden geïdentificeerd kan worden en hoe deze verklaart wordt door de gebouwde omgeving. Daarbij wordt een antwoord gegeven op de volgende twee vraagstellingen: Hoe verhoudt het reisgedrag in suburbane gebieden zich tot dat in stedelijke en landelijke gebieden? En wat is de invloed van verschillende gebouwde omgevingskenmerken op het reisgedrag in deze gebieden?

Dit onderzoek is uitgevoerd met een case studie in de provincie Zuid-Holland. Omdat de definitie van suburbane gebieden niet eenduidig is, zijn er met behulp van diverse indicatoren gebiedstypen onderscheiden. Dit zijn indicatoren zoals afstand tot treinstations, verhouding van banen en huishoudens in een bepaalde straal en autobezit in de wijk.

Voor de gebouwde omgeving is ook een eigen formule ontwikkeld om de indicator "nabijheid van voorzieningen" te berekenen. Deze formule berekend een nabijheidsscore voor de postcodes in Zuid-Holland. Deze formule geeft een gewogen score op basis van verschillende voorzieningen in een bepaalde straal tot het punt waarop de behoefte aan extra voorzieningen van het specifieke type minder wordt, een concept dat bekend staat als marginale utiliteit. Bijvoorbeeld, als er meer supermarkten in de buurt zijn, wordt de toegevoegde waarde van een extra supermarkt kleiner.

Vervolgens is het reisgedrag per gebiedstype (Stedelijk, suburbaan, landelijk) geanalyseerd. Daarnaast is er op basis van omgevings- en sociaaleconomische kenmerken (leeftijd, huishoudgrootte, autobezit en inkomen) een kans berekend om bepaald reisgedrag te vertonen.

De methode van dit onderzoek bestaat uit drie onderdelen, eerst een clustering, gevolgd door een verklaring van clusterlidmaatschap. Voor de clustering is gebruikgemaakt van een niet vaak gebruikte tweevoudige Latente klasse Clusteranalyse (LCCA). Hierbij wordt enerzijds de gebouwde omgeving en anderzijds het vertoonde reisgedrag geclusterd. Als derde stap worden deze twee clusters met elkaar vergeleken.

De LCCA resulteert in 7 clusters voor zowel de gebouwde omgeving als reisgedrag. Met behulp van deze clustering zijn analyses uitgevoerd waar uit blijkt dat het autogebruik in suburbane gebieden tussen dat van stedelijke en landelijke gebieden ligt. Op korte afstanden is het autogebruik in suburbane gebieden tot wel 3 keer hoger vergeleken met stedelijke clusters, zie Table 1. Wel is te zien dat er variatie bestaat tussen verschillende gebieden die als suburbaan worden aangeduid. Verder is zichtbaar dat de keuze voor actieve vervoerswijze zoals fietsen of lopen in elk gebiedstype wordt gemaakt, hoewel het gebruik lager ligt in landelijke en suburbane gebieden. Verder spelen factoren zoals leeftijd, individueel autobezit en nabijheid van voorzieningen een rol in kans op bepaald reisge-

drag. Dit is aangetoond met behulp van verschillende persoonsprofielen en de daarbij behorende kans op het te vertonen reisgedrag. Hierin is te zien dat een hoog autobezit met een laag aantal voorzieningen leidt tot een hoge kans op autogebruik, terwijl een laag autobezit, nabije ligging tot stations en nabijheid van voorzieningen leidt tot een lage kans op autogebruik.

Aantal trips Van en naar het werk met Auto	Suburbaan Cluster 1.0	(Stedelijk) /Suburbaan Cluster 2.0	Landelijk /Suburbaan Cluster 3.0	Landelijk Cluster 4.0	Landelijk Cluster 5.0	Stedelijk Cluster 6.0	(Stedelijk) /Suburbaan Cluster 7.0
0 - 2 km	18.3% (n = 105)	8.8% (n = 32)	15.9% (n = 52)	15.9% (n = 21)	19.1% (n = 26)	5.2% (n = 14)	13.4% (n = 19)
2 - 5 km	36.1% (n = 324)	18.9% (n = 119)	41.7% (n = 130)	52.1% (n = 49)	37.2% (n = 42)	11.0% (n = 46)	23.4% (n = 71)
5 - 15 km	55.2% (n = 907)	43.9% (n = 311)	62.5% (n = 423)	74.9% (n = 137)	62.4% (n = 239)	34.6% (n = 143)	34.0% (n = 150)
Totaal	55.8% (n = 2995)	38.3% (n = 1152)	58.5% (n = 1348)	69.5% (n = 631)	65.2% (n = 749)	24.4% (n = 466)	37.9% (n = 505)

Table 1: Aandeel autogebruik per afstandsklasse. Voor afstanden van 0 tot 2 kilometer wordt 18,3% van de verplaatsingen in suburbane gebieden met de auto afgelegd.

Uit het onderzoek blijkt dat de hypothese dat het autobezit en gebruik hoger ligt in suburbane gebieden klopt. Ondanks een redelijk voorzieningsniveau en de beschikbaarheid van alternatieven zoals fietsen en OV, kan geconcludeerd worden dat het autogebruik in suburbane gebieden opvallend hoog en soms zelfs vergelijkbaar is met dat in landelijke gebieden. Het onderzoek verschaft daarnaast specifieke inzichten over de verschillen in reisgedrag tussen de gebieden in Zuid-Holland, zoals invloeden van de bebouwde omgeving op deze verschillen en het verschil in reisgedrag voor woon-werkverkeer.

Voor toekomstig onderzoek zijn er enkele aanbevelingen om de kwaliteit van de analyses te verbeteren. Ten eerste zou het waardevol zijn om een grotere steekproef te nemen, zodat de gegevens van één jaar voldoende zijn voor representatieve resultaten. Daarnaast is een betere vertegenwoordiging van de Zuid-Hollandse bevolking wenselijk, aangezien onder andere lagere inkomens momenteel ondervertegenwoordigd zijn. Tot slot zou het nuttig zijn om de reisgedragdata op een fijnmaziger schaalniveau, zoals buurtniveau in plaats van wijkniveau, te analyseren. Hierdoor kunnen ook factoren, zoals parkeerbeleid, beter worden meegenomen. In dit onderzoek was dit niet mogelijk, omdat de woonlocatie van respondenten in ODiN op wijkniveau zijn vastgelegd.

Voor toekomstig beleid biedt dit onderzoek waardevolle inzichten en mogelijkheden. Zo geeft dit onderzoek en een op dit onderzoek gebaseerd dashboard[1], informatie over reisgedrag in verschillende typen gebieden. Dit kan helpen in het identificeren van gebieden waar een mobiliteitsbeleid effectief kan zijn. De schattingen van clusterlidmaatschap kunnen inzicht geven in welk type reisgedrag verwacht kan worden, zodat ruimtelijke ontwikkelingen hierop gestuurd kunnen worden.

Voor toekomstig beleid wordt vanuit deze inzichten aanbevolen om te sturen in ontwikkelingen rond suburbane treinstations zowel in het type woningen als specifiek mobiliteitsbeleid zoals het koppelen van een verplicht OV-abonnement aan nieuwe woningen in suburbane stationsgebieden. Ook is een aanbeveling om meer onderzoek uit te voeren in de redenen achter het hogere autogebruik in suburbane gebieden. Daarnaast moet er ook gezocht worden naar samenwerkingen met het bedrijfsleven om duurzame mobiliteit te stimuleren.

Dit onderzoek toont verschillende clusters van de gebouwde omgeving aan. De vergelijking van reisgedrag tussen deze clusters toont aan dat er hierin veel verschillen bestaan. Voor sommige aftandsklassen, waaronder korte afstanden, is het autogebruik in suburbane gebieden tot wel 3 keer hoger vergeleken met stedelijke gebieden. De resultaten van deze analyse kunnen gebruikt worden voor verschillende beleidstoepassingen. Vervolgonderzoek kan aanvullende inzichten opleveren met betrekking tot dit onderwerp.

[1]Zie dashboard op deze website: maxkl.nl, Voor toegangstoken naar reisgedragdata mail m.r.klootwijk@gmail.com

Summary

The province of Zuid-Holland has experienced a large increase in both car ownership and use over the past 20 years. This creates several challenges, such as traffic congestion, high space requirements for cars and challenges in terms of future housing development. Car ownership is expected to further increase in the future.

The increase in car ownership is mainly visible in areas outside cities and therefore these areas deserve specific attention. Those areas can be divided into several subcategories, such as rural areas, but also areas in-between the city and rural areas. The last category, also referred to as suburban area, needs more in-depth analysis because of its increase in car ownership last 20 years. This increase is remarkable, as suburban areas do have proximity to the city, where plenty of amenities are available, and in addition, suburban areas also have alternatives to car use, such as public transport and cycling.

The aim of this study is to investigate suburban areas in more detail. The main research question is which type of travel behaviour can be identified in suburban areas and how it is explained by the built environment? This will answer the following two questions:
How does travel behaviour in suburban areas relate to urban and rural areas?
How do different built environment characteristics affect travel behaviour in these areas?

This research was conducted by a case study in the province of Zuid-Holland. As the definition of suburban areas is not uniform, various indicators were used to identify area types. These are indicators such as distance to train stations, job / household ratio in a given radius and car ownership in the neighbourhood.

For the built environment, a self developed formula is introduced to calculate the ‘proximity of facilities’ indicator. This formula calculates a proximity score for postcodes in the province of Zuid-Holland. This formula gives a weighted score based on different facility types in a given radius up to the point where the need for additional facilities becomes less, a concept known as marginal utility. For example, if there are more supermarkets nearby, the added value of an additional supermarket becomes less.

Subsequently, travel behaviour was analysed for each area type (urban, suburban, rural). In addition, based on built environment and socio-economic characteristics (age, household size, car ownership and income), the probability of showing certain travel behaviour was calculated.

The methodology of this study consists of three parts, first clustering, followed by an explanation of cluster membership. For clustering, a Latent Class Cluster Analysis (LCCA) was used. This involves clustering the built environment and the travel behaviour. As a third step, the clusters of the built environment and travel behaviour are compared.

The LCCA results in 7 clusters for both the built environment and travel behaviour. Using the clusters, analyses were carried out revealing that car use in suburban areas is in-between the urban and rural areas. For short distances, car use in suburban areas is up to 3 times higher compared to urban clusters, see Table 2. However, it can be seen that there is variation between different areas classified as suburban. It is also visible that the choice of active modes such as cycling or walking is made in every area type, although use is lower in rural and suburban areas. Furthermore, factors such as age, individual car ownership and proximity to facilities play a role in likelihood of certain travel behaviour. This was demonstrated using different person profiles and the corresponding probability of the travel behaviour to be exhibited. This shows that high car ownership with a low number of facilities leads to a high probability of car use, while low car ownership, proximity to stations and proximity to facilities leads to a low probability of car use.

Number of trips from and to work with car	Suburban Cluster 1.0	(City) /Suburban Cluster 2.0	Rural /Suburban Cluster 3.0	Rural Cluster 4.0	Rural Cluster 5.0	City Cluster 6.0	(City) /suburban Cluster 7.0
0 - 2 km	18.3% (n = 105)	8.8% (n = 32)	15.9% (n = 52)	15.9% (n = 21)	19.1% (n = 26)	5.2% (n = 14)	13.4% (n = 19)
2 - 5 km	36.1% (n = 324)	18.9% (n = 119)	41.7% (n = 130)	52.1% (n = 49)	37.2% (n = 42)	11.0% (n = 46)	23.4% (n = 71)
5 - 15 km	55.2% (n = 907)	43.9% (n = 311)	62.5% (n = 423)	74.9% (n = 137)	62.4% (n = 239)	34.6% (n = 143)	34.0% (n = 150)
Totaal	55.8% (n = 2995)	38.3% (n = 1152)	58.5% (n = 1348)	69.5% (n = 631)	65.2% (n = 749)	24.4% (n = 466)	37.9% (n = 505)

Table 2: Share of car use by distance class. For distances from 0 to 2 kilometres, 18.3% of trips in suburban area of cluster 1 are made by car.

The study shows that the hypothesis that car ownership and use is higher in suburban areas is correct. Despite a reasonable level of facilities and the availability of alternatives such as cycling and public transport, it may be concluded that car use in suburban areas is remarkably high and sometimes even comparable to rural areas. In addition, the study provides specific insights on the differences in travel behaviour between areas in Zuid-Holland such as influences of the built environment on these differences and the difference in travel behaviour for commuting.

For future research, there are some recommendations to improve the quality of the analyses. First, it would be preferable to have a larger sample, which ensures that one year's data is sufficient for representative results. In addition, better representation of the population of Zuid-Holland is desirable, as lower-income households are currently under-represented. Finally, it would be useful to analyse travel behaviour data at a finer scale, such as neighbourhood level instead of district level. This would also allow for better inclusion of factors such as parking policies. In this study, this was not possible, as the residential location of respondents in ODiN was recorded at the district level.

For future policy, this research provides valuable insights and opportunities. For instance, this research and a dashboard based on this research[2], information on travel behaviour in different types of areas. This can help in identifying areas where mobility policies can be effective. The calculated probabilities of travel behaviour can provide insight into what type of travel behaviour can be expected, allowing spatial developments to be guided accordingly.

Based on these insights, it is recommended to guide developments around suburban train stations, both in terms of the type of housing and specific mobility policies, such as linking a compulsory PT-card to new houses in suburban station areas. Additionally, it is advised to conduct further research into the reasons behind higher car use in suburban areas. Furthermore, collaborations with the private sector should be explored in order to promote sustainable mobility.

This study shows different clusters of the built environment. The comparison of travel behaviour across these clusters shows that there are many differences among the clusters. The results of this analysis could be used for various policy applications. Follow-up research can provide additional insights on this topic.

[2]See the dashboard on: maxkl.nl, For an access token to travel behaviour data, email m.r.klootwijk@gmail.com

Table of Contents

[title=Table of Contents]

List of Figures

List of Tables

List of abbreviations

Abbreviation	Meaning
BE	Built Environment
TB	Travel Behaviour
LCCA	Latent Class Cluster Analysis
LCA	Latent Class Analysis
BIC	Bayesian Information Criterion
BVR	Bivariate Residual
ODiN	Dutch Travel Behaviour Survey
PT	Public Transport
HH	Household
MRDH	Metropolitan Region Rotterdam The Hague
CBS	Dutch Statistics
OAD	Address density
RQ	Research Question
LL	Log-likelihood

Table 3: Abbreviations used in the thesis

1

Introduction

1.1. Problem definition

Over the past decades, the province of Zuid-Holland has experienced a significant increase in car ownership (Provincie Zuid-Holland, 2023c). This increase causes issues like congestion, challenges in housing developments, use of limited public space for parking and a decline in the availability of public transport. Suburban municipalities experienced a high increase in car ownership, sometimes comparable to rural areas. This is surprising, given fact that these municipalities have a closer proximity to facilities compared to rural areas. These suburban areas contribute substantially to traffic problems due to their transport relationship with urban cores in Zuid-Holland. In the future, the negative effects of rising car ownership are expected to intensify, as car ownership is projected to further increase (KiM, 2022).

Figure 1.1 shows that some municipalities had an car ownership increase of 25% or more. Some built environments of these municipalities could be classified as suburban, which means that the neighbourhoods are neither rural or fully urban. The trend of increasing car ownership is noticeable, considering that local and regional governments have actively promoted use of active transportation modes and public transport for several years.

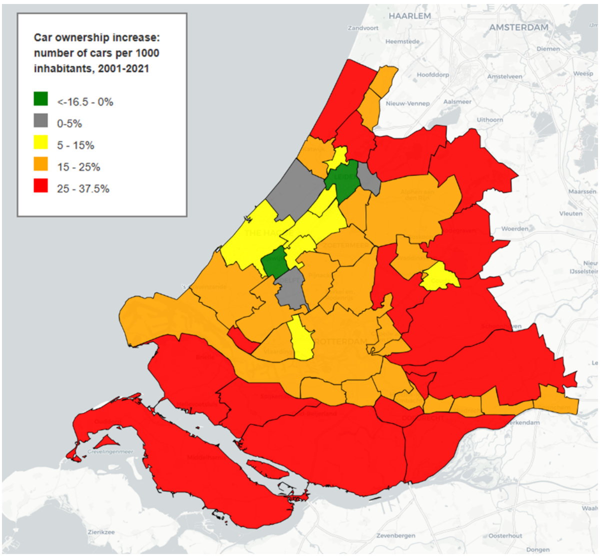


Figure 1.1: Car ownership increase per 1000 inhabitants, growth between 2001 and 2021 (Provincie Zuid-Holland, 2023c).

.

This policy is reflected in recent coalition agreements (Provincie Zuid-Holland, 2011, 2015, 2023d) where the coalition partners intended to promote walking, cycling and public transport (PT) usage among residents. Examples of the policy implementations are prioritizing active modes and PT in new developments, minimizing fare increases in PT and developing new PT connections. A lot of local governments share the same ambitions to promote active modes and public transport.

Despite policies promoting non-car modes of transport, car ownership has increased. Particularly in suburban and rural areas of Zuid-Holland (Provincie Zuid-Holland, 2023c). A province which could be described as a polycentric urban region, with a mix of suburban and rural areas in-between. Most areas in the province are connected by national and provincial roads as well as public transport connections. Since provincial roads (N-roads) are mainly located in suburban and rural areas, the increase in car ownership in these regions is likely result in higher usage of these roads. This is relevant for the province, as it is responsible for daily operation and maintenance of these roads.

The increase in car ownership causes several issues such as increased traffic congestion, challenges for housing developments, use of limited public space and decline in the availability of public transport.

Housing developments for instance, are complicated due to increased car use, as additional traffic generated by these housing developments can create significant congestion issues. While some traffic-related challenges can be mitigated, certain housing developments may require modification in number of houses due to their expected impact on mobility. So increasing car usage poses challenges to the ambitious housing development goals in the province of Zuid-Holland aiming to build 240.000 houses (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2022). Moreover, cars also need to be parked in public space. This public space could also be used for new/more houses or urban green spaces. In addition, due to increased car use, the usage of public transport is decreasing. This results in higher operating costs per passenger, further weakening the financial profitability of public transport services. which is particularly relevant for public transport in rural areas. This are currently areas with a higher subsidy rate on PT (Autoriteit Consument en Markt, 2021). In rural areas, it is visible that highly subsidised, less profitable PT services disappear. This actually increases car use/ownership for people living there.

1.2. Research gap

Suburban and rural areas contribute substantially to the rise in car ownership in Zuid-Holland (Figure 1.1). The traffic generated in suburban areas is primarily concentrated between suburban and urban regions. This makes it interesting to conduct comparisons between different types of built environment areas, however there is no clear definition of a suburban area (Boomkens et al., 1997; Forsyth, 2012).

Literature on the relationship between travel behaviour and the built environment partly describes how suburban areas influence travel behaviour in the Dutch and international context. Ewing and Cervero, 2010; Stead and Marshall, 2001 conducted literature reviews on this topic, highlighting the relationship between the built environment and travel behaviour. This reveals several relationships between mode choice and built environment. However these studies are limited due to their focus on urban areas, with only a few studies addressing suburban and rural areas. The suburban areas are important due to their interaction with the city, i.e. a lot of trips are from suburbs to the city. Among the studies that focus on suburban areas, the focus is primarily on travel behaviour-related data (Chen & McKnight, 2007; Chen et al., 2007; Friedman et al., 1994). In these studies it is visible that suburban areas demonstrate a higher use of car and longer trip distances compared to other areas.

In addition to the limited focus on travel behaviour in suburban areas, characterised as the intermediate phase between urban and rural areas (Boomkens et al., 1997; Forsyth, 2012), most studies on travel behaviour in suburban areas are conducted in the United States. These studies reveal different patterns of travel behaviour compared to the Netherlands also other European studies show a significantly higher use of active modes of transport (Friedman et al., 1994; Liu & Alain, 2014). Moreover, in the Netherlands the widespread use of bicycles results in unique travel patterns compared to travel behaviour in other European regions (Pucher & Buehler, 2017).

The historical development of cities differs between Europe, the United States and other urban agglomerations. In Europe, this resulted in a different definition and appearance of suburban areas. However even within European countries there are differences between urban agglomerations. For example Zuid-Holland has an urban structure that is uncommon in the United States and other European countries (GHSL - Global Human Settlement Layer, 2016). In fact, the province of Zuid-Holland can be described as a collection of cities with suburban area's in-between, leading to polycentric travel patterns (Boussauw et al., 2012; Planbureau voor de Leefomgeving, 2006). A similar European urban agglomeration is the Manchester, Liverpool and the Leeds area (Dembski, 2015).

It could be seen that there are various types of suburban areas, distant suburban areas or suburban areas adjacent to the city. Sometimes this is resulting from the aforementioned historical urban developments (Kubeš & Ouředníček, 2022). However, when a suburban area is defined it is often seen as just one type of area. There is no clear literature that examines the relationships between different types of suburban areas and travel behaviour. While this is interesting for analysis of specific travel behaviour as well as built environment indicators influencing the travel behaviour.

As described above, differences across specific built environment areas like suburbs, are often neglected in the current literature, possibly due to challenges in defining whether an area is urban, suburban or rural. This challenge creates difficulties in understanding the relationship between various travel patterns and the influence of the built environment in suburban area's. This is even further complicated by the fact that there are different types of suburban areas. The role of these suburban areas on travel behaviour remains unclear.

In conclusion, literature provides knowledge on the influence of general mechanisms on travel behaviour. The main gap is the unclear role of various types of suburbs on travel behaviour. Therefore it is interesting to identify different area types and conduct travel behaviour comparisons across different area types in order to improve policymaking for 'suburban' areas with a high car ownership increase, despite proximity to facilities or urban cores,

1.3. Research questions

To address this knowledge gap and find the relationship between the built environment and travel behaviour in suburban areas, research questions have been formulated. The main research question is formulated as follows:

Main RQ: What types of travel behaviour can be identified in suburban areas and to what extent are they explained by the built environment?

To support the main research question the following sub-research questions are formulated:

- **Sub-RQ1:** What clusters can be identified on travel behaviour?
- **Sub-RQ2:** What clusters can be identified on the built environment?
- **Sub-RQ3:** What is the impact of built environment on clusters of mobility patterns?
- **Sub-RQ4:** How does travel behaviour differ in different suburban types in Zuid-Holland?

1.4. Research approach

Given the patterns in travel behaviour data and the built environment defined as area type based on multiple indicators, Latent Class Cluster Analysis (LCCA) was employed to identify subgroups within the data. LCCA allows to identify hidden subgroups within the data, each may have unique travel behaviour or distinct built environment characteristics. Applying LCCA, gives better understanding of the complex relationships between travel behaviour and the built environment clusters across different segments of the population.

The specific method for clustering is LCCA as it outperforms other clustering methods like K-means

(Magidson & Vermunt, 2002). LCCA is a useful method that identifies clusters which are homogeneous within the cluster and as heterogeneous as possible across clusters.

Using datasets from the built environment (2020, 2022) and travel behaviour (2018,2019 and 2023), containing approximately 30.000 individuals and 512 postcodes, it is possible to make a built environment and travel behaviour cluster model. The two estimated cluster models could be linked trough the postcode which serves as a connection between the built environment and individual travel behaviour clusters. This approach allows for analysis on the relationships between these clusters and different covariates.

1.5. Thesis outline

This report is structured as follows: First, the literature section reviews the current literature on the thesis topic and identifies a knowledge gap. Accordingly followed by the data and methodology section to address the research questions chapter 3. The results from the study are then presented in chapter 4, followed by chapter 5 with the conclusion, the discussion in chapter 6 and recommendations (chapter 7) for future research and policy implementation.

2 Literature

This section further expands on the topics from the introduction, with the aim of developing a conceptual model to explore certain relationships. The focus of this literature section lies on suburban areas and urban development followed by a discussion of current knowledge of travel behaviour and the built environment. Various relationships between the built environment and travel behaviour are analysed, such as the influence of proximity to facilities. Finally, this section outlines what is known from the literature and what knowledge is still missing.

The literature was collected using 3 methods (vom Brocke et al., 2015). The first method involved the use of specific search terms (Bramer et al., 2018; James Cook University, 2023), the second and third methods include forward and backward snowballing, respectively. Forward snowballing looks at studies that cite a the article, whereas backward snowballing reviews the references within the article (Jalali & Wohlin, 2012). Forward and backward snowballing provide the opportunity to discover literature that might otherwise be overlooked. Combining these methods allows for the discovery of appropriate and relevant literature.

2.1. Suburban areas and urban development

Various terminologies are used to define the types of built environments. For example, CBS categorises areas into 5 area types, ranging from highly urban to non-urban, which is based on address density (CBS, 2024b). Others classifications have a more qualitative approach (ruimtemettoekomst.nl, 2013).

Internationally, three types of built environments are mostly distinguished: urban, suburban and rural areas (Forsyth, 2012; Harris & Larkham, 2004). However, there is no uniform definition of suburban areas (Forsyth, 2012). Hamers, 2003 defines suburban areas by what they are not, they are neither urban or rural. Furthermore suburban areas show variation across themselves, some are more rural-suburban while others show a more urban-suburban character (Boomkens et al., 1997).

Besides the challenge in definition, the history of suburban areas also plays an important role. For instance, Dutch areas that can be defined as suburb have a diverse history. In the past there have been 'Groeikernen', 'Bloemkoolwijken' (cauliflower neighbourhoods), Vinex neighbourhoods and Transit-Oriented Developments (TOD) in the Netherlands. These developments are the result of specific policy measures (Abrahamse, 2019). Due to the unique character, these developments are not always directly comparable to other countries.

In the US, urban development of some areas experienced a phenomenon known 'urban sprawl', characterised by residential areas only having residential functions without other amenities. This low function mix leads to a increased need to travel long distances, often by car. In response a counter movement started to emerge, called 'smart growth' or 'traditional neighbourhood development' (TND). Smart growth and TND prioritize a higher function mix in the neighbourhood leading to less car use and shorter travel distances (Cervero & Radisch, 1996; Khattak & Rodriguez, 2004).

The lack of a clear definition of suburban area combined with the fact that there are different types of suburban area (with possibly different travel behaviour), complicates the identification of suburban. Especially since there are international variations in suburban area characteristics. This poses challenges for conducting research on suburban areas. (Hamers, 2003). Therefore, it is desirable to analytically identify different types of suburban areas.

2.2. Travel behaviour and suburban areas

In terms of travel behaviour in suburban areas, differences are notable. Internationally, suburban areas are generally characterised by higher car usage compared to cities, while the use of active modes of transport, such as walking and cycling are lower (Friedman et al., 1994; Liu & Alain, 2014). In addition, studies indicate a correlation between distance from a major urban centre and car use: the further an area is from the urban centre, the higher the car usage (Dandy, 1997; Vega & Reynolds-Feighan, 2009a).

Some of the studies mentioned earlier were not conducted in the Netherlands, so specific Dutch context is missing. In the Netherlands, unique factors such as high bicycle use (Pucher & Buehler, 2017), the compact urban form of Zuid-Holland Holland with its close city centres (Batty, 2001; GHSL - Global Human Settlement Layer, 2016) and the polycentric travel behaviour characteristics of the Randstad region(Boussauw et al., 2012; Planbureau voor de Leefomgeving, 2006) influence travel behaviour. This may lead to different travel behaviour making it difficult to apply findings from international studies directly to the Dutch context. Moreover, there is a lack of suitable literature available that directly compares different area types and travel behaviour within the Netherlands. Most studies focus on a single mode of transport, such as car use, or on specific methods of characterizing areas, such as address density. While a focus on suburban areas is actually interesting.

The challenge of defining suburban areas uniformly, combined with the lack of knowledge about travel behaviour in different types of suburban areas makes it interesting to differentiate between several area types and compare them with the existing travel behaviour.

2.3. Relation travel behaviour and built environment

The built environment is often defined in various ways as done by (Brownson et al., 2009; Frank & Engelke, 2001; Handy et al., 2002; Hassler & Kohler, 2014). A widely cited article by (Handy et al., 2002) defines the built environment using a combination of the terms urban design, land use and the transportation system.

”Urban design” refers to the physical design of the built environment, including the function and appearance of public spaces. ”Land use” encompasses the types of activities that occur in the built environment and includes concepts such as density. The third term, ”transportation system,” covers infrastructure like roads as well as the service levels of transportation options, such as distance to train stations. The built environment is defined as a combination of three elements; urban design, land use and the transportation stystem (Brownson et al., 2009; Handy et al., 2002). Within the built environment, various activities occur, resulting in the individual need for transportation.

Some research describe that the mode choice is based on reaching maximum utility within budget constraints(Crane, 1995), Boarnet and Crane, 2001 expanded on this, by arguing that travel costs are influenced by built environment, while others advocate for attitudes and sociodemograpic factors influencing travel behaviour (Lenormand et al., 2014; Ton et al., 2020). Literature reviews (Ewing & Cervero, 2010; Stead & Marshall, 2001) introduce different built environment factors as variable in the travel behaviour and mode choice. In this field, it can be seen that several factors influence travel behaviour.

Address density is often used as an indicator of the built environment (Cervero & Radisch, 1996; Kockelman, 1997). However, this indicator does not always provide a complete representation. For example, in the Netherlands, some areas are classified as urban based on address density, have facilities located at a considerable distance. (Martensen & Arendsen, 2024b), Therefore, it is desirable to develop an alternative indicator that combines density with proximity to facilities.

The land use mix, job/housing balance is regularly used to define the built environment. A high job-housing mix can create opportunities for work close to the residential location, resulting in short travel times for residents (Cervero & Duncan, 2006; Stoker & Ewing, 2014). However, this effect applies in a greater extent to more divergent balance ratios, according to Peng, 1997.

The distance or travel time to a transit stop is also regularly included as a factor to examine the built environment (Ewing & Cervero, 2010). This factor serves as an explanatory factor in choice to travel behaviour (Krygsman et al., 2004; Lunke & Engebretsen, 2023).

Car ownership in a neighbourhood has also been used as an indicator in this study, similar to previous research by Martensen and Arendsen, 2024a. Where car ownership serves as a class indicator. The number of parking spaces influences the travel behaviour of individuals in the neighbourhood (P. Christiansen et al., 2017; Gruyter et al., 2024). In individual car ownership there is a difference between area types, although this variation has not been extensively documented in the Dutch context in current literature (Ewing & Cervero, 2001; Gruyter et al., 2024). Car ownership influences the travel behaviour, individuals owning a car are using the car for more trips compared to those who do not own a car (Dieleman et al., 2002; Van Acker & Witlox, 2009).

Various social variables influence travel behaviour. Variables such as age, household size, car ownership and income have impact on travel behaviour (Dieleman et al., 2002; Stead & Marshall, 2001; Van Acker & Witlox, 2010). For example a higher age is associated with fewer (car)trips and shorter trip distances (Schwanen et al., 2004; Van Acker & Witlox, 2010). Larger households tend to make more car trips on household level although less trips per person (Son et al., 2013). Car ownership correlates with more car trips and longer distances travelled (Van Acker & Witlox, 2009). Income also shows a positive effect on the number of car trips while negatively impacting PT and active mode trips (Dargay & Hanly, 2004).

Given the impact of social variables and built environment categories on travel behaviour choices, it is important to include these factors in the study alongside the comparison of travel behaviour and built environments. Specifically, investigating how these variables influence different types of built environments is interesting.

2.4. Conclusion and conceptual model

The definition of suburban areas is not uniform, in addition, there is some variation among areas defined as suburban section 2.1. This diversity makes it interesting to identify different types of suburban areas and analyse to what extent they differ in terms of travel behaviour.

As previously mentioned, comparing travel behaviour with built environment characteristics provides meaningful insights. Furthermore, examining the relationships between social factors and Built environment variables can increase our understanding of travel behaviour. These insights could contribute to mobility policies in the Dutch and Zuid-Holland context.

A conceptual model has been developed to examine the influence of built environment clusters on travel behaviour clusters, as illustrated in Figure 2.1.

The conceptual model consists of three main components: (1) clusters of the built environment, based on BE indicators, (2) clusters of travel behaviour, based on trips made, (3) set of variables. These main components are linked in this conceptual model. First, given the built environment the probability of certain travel behaviour is examined. So for each built environment cluster the travel behaviour probability is examined. This provides insight in the differences in travel behaviour. Secondly each built environment cluster is described based on socio-demographic variables and built environment characteristics. This includes identifying the types of people living in a particular built environment type and built environment characteristics in each built environment cluster. A similar approach is applied to describe the Travel behaviour clusters.

It is expected that rural cluster show a more car oriented travel pattern compared to the urban clusters. In terms of travel behaviour, suburban clusters are expected to be in-between rural and urban clusters.

This conceptual model forms the basis for this research. chapter 3 elaborates on the categorization of BE and TB clusters, as well as specific method used to analyse the influence of BE and socio-demographic variables on the probability of TB cluster membership.

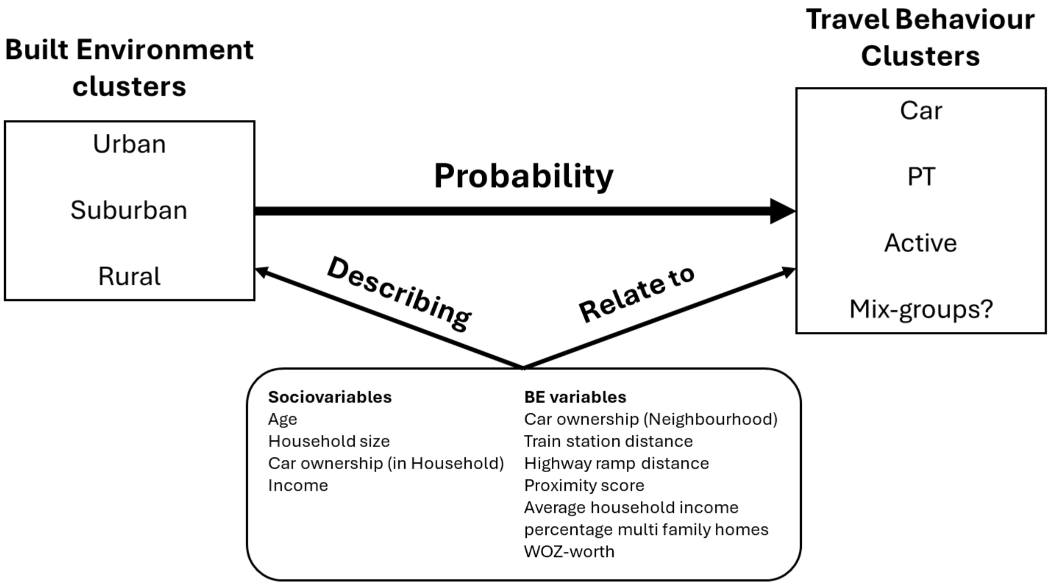


Figure 2.1: Conceptual model

3

Methodology

This section outlines how the research questions are addressed. First, the study area is explained, followed by a description of the data and relevant data processing steps. Then, the methods applied to analyse the data are discussed. By the end of this section, it becomes clear how the research was conducted.

3.1. Study Area

The study area of this thesis is the province of Zuid-Holland, the most densely populated province in the Netherlands. This region is chosen due to its polycentric mobility behaviour (Boussauw et al., 2012), large cities such as The Hague and Rotterdam as well as suburbs and rural area's in between. For analysis such as proximity to facilities, data beyond the provincial boundaries is considered, if available in order to avoid inaccurate measurements and unequal weighting of the study area edges.

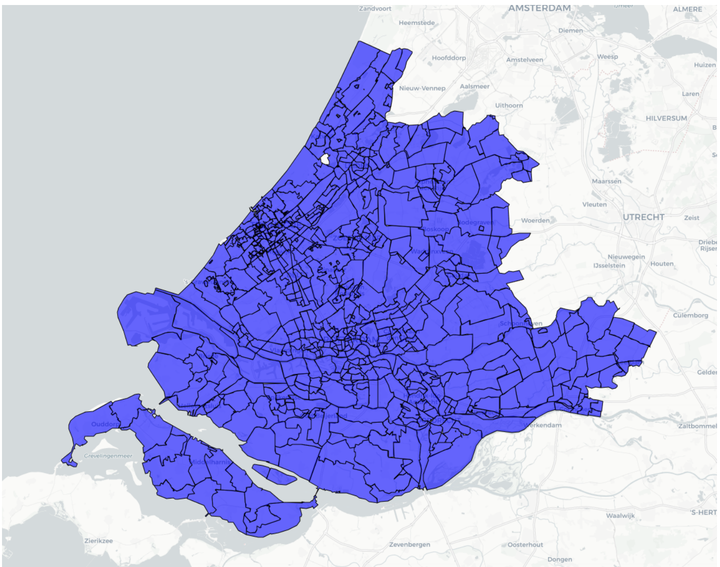


Figure 3.1: Study area, the province of Zuid-Holland visualised on a map.

3.2. Research data

This section outlines the data selection and processing methodology. First the built environment data, is described, with the introduction of new formulas to calculate job-household balance and proximity. Next to the built environment data, the method for gathering and processing the travel behaviour data is elaborated. The travel behaviour data is reviewed based on the real population in Zuid-Holland.

3.2.1. Built environment Data

Built environment data is collected at the Dutch PC4 (Postcode-4) level (1234 xx) to ensure comparability with the travel behaviour data, where respondents locations are specified at the PC4 level.

Built environment data, including the accessibility to facilities, population demographics and household characteristics originates from CBS PC4 data (CBS, 2020a, 2022a). This data contains information related to the postcodes. Job data is obtained from processed LISA data (LISA, 2022) LISA is a database with all addresses of jobs in the Netherlands. Car ownership data is also obtained from the CBS (CBS, 2023b)

Postcodes with fewer than 100 households are excluded due to CBS data limitations, as data for areas with fewer than 100 households is anonymised. Additionally, this threshold minimises the influence of outliers making the data more reliable. After these exclusions, 512 from the 537 postcodes in Zuid-Holland are included in the analysis. Indicators for the built environment data are visualized in Table 3.1.

Indicators and covariates(n = 512)	**Mean**	**Median**	**SD**	**Minimum**	**Maximum**
OAD	2260	1810	1847	56	8.637
Median construction year	1972	1970	19	1940	2020
Percentage multi-family homes	42,1%	33,5%	29,75%	0%	100%
Mean Household-income	€ 34.070	€ 33.700	€ 7,84	€ 18.200	€ 120.900
WOZ-worth	€ 334.037	€ 322.000	€ 116.590	€ 161.000	€ 1.448.000
Train-station Distance (km)	6.6	3.9	7.41	0.5	42.6
Highway ramp (km)	1.87	1.6	1.07	0.2	6.4
Car ownership per household	0.92	0.97	0.27	0.25	1.6

Table 3.1: Built environment Postcode-4 data (CBS, 2020a, 2022a). Bold text represents the indicators, rest are covariates

Proximity score
It is found that the OAD (Address density) only counts the addresses and therefore not always represent the availability of facilities in a neighbourhood such as hospitals, shops, supermarkets or restaurants, this results in misinterpretations of some neighbourhoods. Some built environments have a high OAD but a lack of facilities particularly in more residential areas like Vinex-wijken[1] and growth cities such as Zoetermeer. These areas show a high OAD but fewer facilities like shops, restaurants and leisure. Other places such as towns in rural areas appear to have sufficient facilities despite a lower OAD. As a result an alternative measure is needed to asses the proximity to facilities.

To get a better representation of proximity to facilities, a proximity score is introduced for this study. This score is calculated using a formula that counts the number of different amenities within a specific range and assigns a weighted value to each amenity. For example, a nearby primary school is given a higher weight than a restaurant. The proximity score is scaled to a median of 100 so 50% of the postcodes is below the score of 100 and 50% above. The formula developed for this research to calculate

[1]Vinex-wijken are new urban developments built between 1995 and +-2005. These neighbourhoods were located near (mid-sized) cities to increase the number of customers for existing stores and to reduce travel distances, to encouraging the use of active transportation modes and public transport.

the proximity score is presented below.

$$Proximityscore = \sum score_i$$

n = number of facilities
lw = number of facilities to start lower weight
median = median number of facilities in Zuid-Holland
i = Facility like supermarket or primary school

$$\begin{aligned} &If(n < lw): \\ &\quad Score_i = \sum \frac{n}{median} \times factor \\ &Else(n > \text{lw}): \\ &\quad Score_i = \sum \frac{n[withinlw] + ((n - valuelw) \times lowerweightfactor)}{median} \times factor \end{aligned} \tag{3.1}$$

The developed formula is tested using dummy variables. To prevent outliers due to a more than necessary number of facilities, a ”countmax” parameter is introduced. This adjustment is introduced to account for the utility contribution of extra similar facilities in a neighbourhood that already has a sufficient facilities. This utility contribution is lower with already a high number of facilities (Banister & Berechman, 2001). For example, the utility contribution of a 10th supermarket is less than the contribution of the first or second supermarket. This concept is called marginal utility (Vandenbergh, 2024).

The minimum number of different facilities to experience freedom of choice is not clearly defined in literature. Therefore, a survey was conducted among 15 mobility experts at the province of Zuid-Holland to determine the minimum number of facilities for perceived freedom of choice. More details and responses to the survey could be found in Table 3.2 and section B.1.

To assign appropriate weights to the facilities, the survey conducted among mobility experts also asked for the most and least important facilities. The weights displayed in Table 3.2 show the adopted weights used in the formula.

The median in the formula was introduced to adjust the score based on the availability of certain facilities. For example, the number of supermarkets is higher compared to the number of hospitals, so dividing by the median helps to correct for this imbalance. In the formula, this is combined with the lower weights (countmax) to account for marginal utility. The combination of the countmax parameter which represents marginal utility and the inclusion of median scores provides reliable results when using dummy variables.

The count radius is based on values from the CBS dataset (CBS, 2022b). In the dataset the 1, 3, 5 and 10 km radius is included for counting the facilities. For each type of facility a range has been selected, based on kind and frequency of use.

A lower weight factor of 0.1 is applied for all facilities, meaning that after the countmax is reached within the defined range, additional facilities contribute only 10% to the proximity score. Using these parameters the proximity score is calculated. Further research is recommended to define the optimal count radius, weight factor, marginal utility of facilities and lower weight factor, see chapter 7.

	Median (number)	Factor	Count radius	Start lower weight (lw)	Lower weight mean from survey (median)
Supermarket	7.1	8	3 km	3	3.9 (3)
Daily groceries	25.9	4	3 km	12	14.1 (11)
Primairy School	1.8	7	1 km	2	2.1 (2)
High School	2.8	4	3 km	3	2.2 (2)
Café	0.8	2	1km	2	1.8 (2)
Cafetaria	2.4	2	1km	2	2.1 (2)
Hotel	3.8	2	5 km	2	2.3 (2)
Restaurant	21.3	2	3 km	4	9.8 (9)
Attraction	3.0	1	10 km	2	2.8 (3)
Cinema	0.9	1	5 km	2	1.6 (2)
Musea	1.1	1	5 km	2	2.7 (2)
Hospital (with beds)	0.9	5	5 km	1	1.5 (1)

Table 3.2: Weights, start of lower weight factors compared to expert survey results. Used in the proximity formula.

Function mix indicator

The function mix indicator indicates the probability of finding a job within a given area and the ratio of inbound to outbound trips. A high ratio means that the number of jobs does not match the number of households so in theory more incoming trips are needed to fullfill all jobs. An average ratio indicates that the number of jobs and households is well-balanced.

The job-household balance, is calculated with Equation 3.2. This is provides insight into the potential for trips within and outside an area. The function mix indicator is calculated at PC4 level and within 3 km radius. The 3 km range level includes all jobs and households within a 3 km distance from each postcode. This score is based on the job-housing balance, which is a commonly used indicator to measure the land-use mix (Cropper et al., 2005; Kockelman, 1997; Spears et al., 2014). Table 3.3 shows the calculated proximity score and function mix indicators are displayed. In Appendix C the values are displayed on a map.

$$\text{job-housing balance}_{pc} = \frac{Jobs_{pc}inrange}{Households_{pc}inrange} \tag{3.2}$$

Indicators and variables (n = 512)	Mean	Median	SD	Minimum	Maximum
OAD	2260	1810	1847	56	8.637
Proximity Score	116.09	100	77.50	2,42	366.54
Jobs per inhabitant in PC4	0.69	0.33	1.74	0.07	24.04
Function mix (Jobs/households) in PC4	1.48	0.74	3.71	0.18	54.32
Function mix range (All jobs/household in a 3km radius from postcode)	1.18	1.12	0.45	0.38	4.58

Table 3.3: Built Environment indicators from own calculations. (bold are indicators) (OAD not from own calculations but included to show difference to proximity score.

An interesting observation from the OAD indicator is the variation in density. The mean and median OAD is relatively high compared to the national average of 1066 addresses in the Netherlands (CBS, 2022a). This confirms the high density in Zuid-Holland. The proximity score also shows some deviation, with the scaling applied to set the median to exactly 100. As illustrated in Appendix C, the proximity to facilities appears logical, high in cities, but also appropriate in rural towns. The jobs function mix indicator shows a high maximum and is mainly concentrated in areas with greenhouse zones, such as the Westland or postcodes with large distribution centres like Bleiswijk. This suggest the job function balance is a good indicator of incoming trip potential, as many trips must be directed to areas with high jobs-to-household balance to meet employment demands.

In addition to the indicators, the dataset includes socio-demographic information and the PC4 codes of the individuals respondents, allowing to link each PC4 from the BE data to Travel behaviour (TB) data. The combination of TB and BE data explained in this section, serves as the foundation for the LCCA and further analysis from the LCCA results.

3.2.2. Travel behaviour data

The used travel behaviour data originates from the annual CBS travel behaviour research/survey “Onderweg in Nederland” (ODiN) (CBS, 2019, 2020b, 2024a). ODiN data contains data of individual travel behaviour within a 24-hour period, including information on travel mode, distance, trip duration, origin and destination as well as sociodemograpic information about the respondents. The ODiN data is selected due to its number of respondents and possibilities of combining data across multiple years. The unprocessed ODiN datasets for Zuid-Holland of 2018, 2019 and 2023 contains information from 41.856 individuals and 145.537 trips.

Although MPN (mobiliteitspanel Nederland) data from KIM, 2021 covers a 72 hour period and offers a more extensive dataset, with the same respondents surveyed annually, the data includes only around 500 respondents in Zuid-Holland. This is around 60 times less compared to the 29.721 respondents in the ODiN data. Given the need for more respondents per postcode to analyse the built environment effectively, ODiN data is more suitable as primary data source for the travel behaviour analysis.

The analysis utilizes data from the years 2018, 2019 and 2023. (CBS, 2019) confirms data from different years can be combined. Specific data selection and exclusion methods to get usefully data insights are more detailed in this section.

Data from 2023 is included to increase the number of respondents per postcode, particularly in rural areas with fewer residents. The 2023 dataset contains post-COVID travel behaviour after the Covid years 2020,2021 and 2022. Differences between 2018/2019 and 2023 include an increase in non-travellers and a decline in public transport usage. The most surprising change is the 5% decrease in total number of trips, particularly PT trips, with a 22% decrease. Further details on the travel behaviour changes trough the years are provided in Table 3.4. The post Covid changes could influence the result of the research. However for the number of data per postcode, the data from 2023 had to be added (2017 and earlier is not available because of a method change). Effects will be mentioned in the discussion.

TB Data SH	**SH 2018 + 2019 (n = 19840)**	**SH 2023 (n = 9881)**
Car trips	1.25	1.12
PT trips	0.23	0.18
Active trips	1.32	1.36
Total trips	2.80	2.66
Distance car	22.96	19.28
Distance total	33.21	28.18

Table 3.4: (After Covid) Differences in mean of ODiN data between 2018 + 2019 and 2023 (Data from the province of Zuid-Holland only)

Data Processing

Some adjustments were made to ensure the data gets more reliable and practically applicable.

The first adjustment is exclusion of additional data on regional level known as ”ODiN meerwerk”. This additional data was performed for the MRDH, a partnership of 23 municipalities in Zuid-Holland, including Rotterdam and The Hague. ’Meerwerk’ was excluded to avoid over-representation of the MRDH region, resulting in the exclusion of approximately 11.000 individuals from the ODiN dataset.

Individuals living in small postcode areas are excluded due to the fact CBS built environment data is anonymized for groups smaller than five persons, due to privacy regulation. As mentioned earlier in section 3.2, the threshold of 100 households per postcode was applied due to data availability and the prevention of outliers. Consequently, 286 individuals were excluded from the dataset.

Trips made during work-time are reported in ODiN data but were excluded from this analysis. As the focus is on individual travel behaviour, some examples of work-time trips could be a package deliverer. As a result 4.688 trips are removed. Individuals who did not make other trips during the day are reported as non-travellers, so no respondents are excluded after this step.

Some trips consist of a series of trips sometimes with different modes, such as a train trip where the bike serves as access mode and walking as egress mode. To maintain clear data, only the primary mode of these serie trips are reported in the travel dataset. So the mode used for the largest distance is included, leading to the removal of 18.052 trips. No respondents were excluded since all trips maintain in the dataset with their primary mode.

All adjustments described above are summarised in Figure 3.2. After these modifications the travel behaviour of 29.721 individuals from the datasets of 2018, 2019, and 2023 is included. Sociodemographic characteristics of these individuals are presented in Table 3.8.

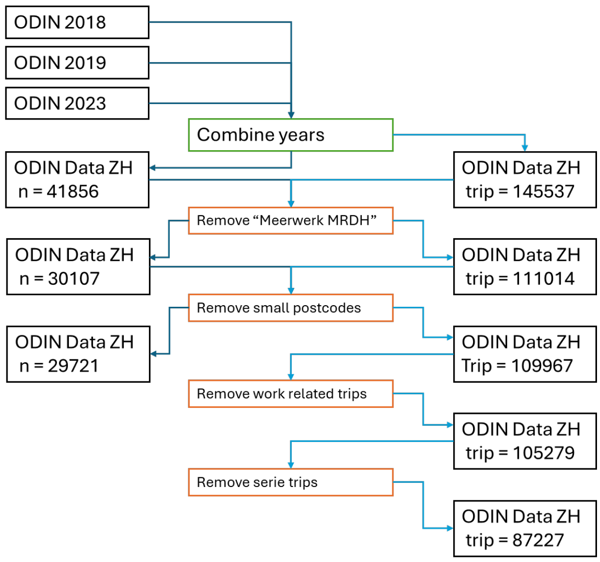


Figure 3.2: Data processing steps, number of respondents (n) and number of trips (verpl).

The trip data in ODiN contains 21 different travel modes. For this research and comparability these modes are categorised into four groups; Car, Public transport, Active modes and Other. The assignment of each ODiN mode to its respective group is visible in Table 3.5.

Motorcycles are placed under to the car mode category due to their speed and place on the road which is comparable to cars. Mopeds are grouped with the active mode group because their usage, position on road and lower speed are more comparable with active modes. The other category includes modes such as boats and tractors, which do not fit with one of the other groups.

Since the "other" mode accounts for only 0.05 trips per person per day, the "other" mode is excluded from the dataset. This adjustment effects the data of 193 respondents where 83 respondents did not make any trip other than a trip with the mode categorised as "other". These respondents were therefore excluded from the dataset.

For each mode described earlier, the number of trips and the corresponding travel time and distance is collected. Car trips under 2 kilometres are counted to indicate the potential for reducing short-distance

Mode (in this Report)	Mode (ODiN dataset) NL and EN translation
Car	personenauto, bestelwagen, motor (car, delivery van, motorcycle)
public transport (PT)	Bus, Tram, Metro, Taxi (bus, tram, subway, taxi)
Active	speedpedelec, elektrische fiets, niet elektrische fiets, te voet, skates/step, anders zonder motor (speed pedelec, electric bike, non-electric bike, On foot, Skates/scooter, Other without motor, moped, light moped)
other	other

Table 3.5: Grouping of Travel modes in this report and ODiN travel modes.

car trips. Additionally, the dataset registers the purpose of trips, categorised into mandatory trips (work or school) and non-mandatory trips (visits, hobbies, and sports). The processing of the data includes a step to calculate the percentage of trips made during peak hours. This allows for the identification of individuals contributing to network congestion during peak hours. The travel behaviour data is visualised in Table 3.6.

Indicator (n = 29721)	Mean	Median	SD	Minimum	Maximum
Car trips	1.21	0	1.65	0	16.0
PT trips	0.22	0	0.63	0	5.0
Active trips	1.33	1.0	1.73	0	17.0
Distance car	21.74	0	45.45	0	560.0
Distance PT	5.76	0	24.4	0	517.1
Distance Active	4.01	0	7.69	0	205.9
Distance total	31.53	13.0	49.04	0	560.0
% agenda-setting	26.3	0	38.29	0	100
% Non-agenda-setting	57.32	66.7	44.43	0	100
% Peak hour trip	30.26	25.0	32.07	0	100
% Car less 2km	4.27	0	17.26	0	100

Table 3.6: ODiN Travel Behaviour Data. Trips, distances and some other information reported.

Data review

Representativeness tests on socio-demographic variables (gender, age, household size, household income and car ownership) revealed differences between the dataset and the average of the population in the Province of Zuid-Holland.(CBS, 2023a; Provincie Zuid-Holland, 2023a, 2023b) All socio-demographic variables were found to be significant different compared to the population. The average age is higher in the sample, The household size and car ownership is higher compared to the population average. This could be due to the under-representation of lower income classes in the ODiN data. More information on the sample is presented in Appendix D.

CBS uses stratification on national level to correct for this kind of errors, this CBS stratification values could not be used on regional level (CBS, 2023a). It was decided to not correct for this using the stratification of the CBS in order to maintain unmodified travel behaviour data. The impact of these differences will be addressed in the discussion section.

The TB data in Table 3.6 reveals some interesting findings. Respondents of the ODiN data take an average of 2,74 trips per day, which closely aligns with the Dutch average as well as the averages in the UK and Denmark (CBS, 2024a; Department of transport, 2024; Knudsen, 2014). In Denmark the model split is almost the same compared to the Netherlands (H. Christiansen & Baescu, 2022). However the share of active mode trips in Zuid-Holland is around 10%-point higher compared to the UK. This is primarily due to the high share of bikes in the Netherlands. As a result of higher bike use, the

Indicator (n = 29721)	Mean	Median	SD	Minimum	Maximum
Age	43.84	44.0	22.23	6.0	99.0
Household size	2.78	2.0	1.41	1.0	10.0
Income Class	6.69	7.0	2.69	1.0	10.0
Car ownership per household	1.24	1.0	0.91	0.0	9.0

Table 3.7: Sociodemograpics of all ODiN Respondents. (CBS, 2023a; Provincie Zuid-Holland, 2023a, 2023b)

Sociodemograpic	Sample mean	Population PZH mean	Test of significance
Age (from 6-older)	43.84	43.52	t = 2,469 df = 29721 p = 0,014
Household size	2.78	2.1	t = 83,084 df = 29721 p = 0,000
Income Class (10% class)	6.69	*	t = 76,264 df = 29721 p = 0,000
Male (%)	51.1	49.4	
Female (%)	48.9	50.6	
Car ownership per household	1.24	1.03	t = 40,172 df = 29721 p = 0,000

Table 3.8: Sociodemograpic data (Whole population) and ODiN sample data compared, (PZH income class mean for the population is estimated as the median of 5.5)

distance travelled by active modes is higher in the Netherlands, as the bike covers far greater distances compared to only walking in the UK.

The distances travelled by individuals show substantial variation, as visible trough the standard deviation (SD). This difference is as expected, given that some people travel long distances while others only make short trips. It is notable that the mean distance travelled by car is higher compared to the distance travelled by public transport. The distance travelled by active modes is lower, as this mode includes walking and cycling, which are typically modes for shorter distances.

The percentage of agenda setting trips show more variation across the dataset. This is as expected since the purpose of the individual trips differs a lot per person per day. The peak hour and less than 2km trips align with expectations and could provide valuable insights for further analysis.

3.3. Latent class cluster analysis

The first two sub-research questions aim to identify clusters in the built environment and travel behaviour. The cluster identification is conducted using indicators from the built environment and travel behaviour. After the cluster identification the influence of variables on cluster membership could be calculated.

Given the diverse nature of travel and built environment data, it is necessary to group the data. By applying a cluster analysis, we can better understand the complex relationships in travel behaviour and built environment clusters across different segments of the population. This leads to meaningful insights in clusters.

The identification of clusters could be performed analytically trough a latent class analysis (LCA), a method chosen for its ability to uncover hidden subgroups within a dataset, resulting in a description of these subgroups (Magidson & Vermunt, 2002). In essence words, LCA identifies individuals who are homogeneous within clusters and heterogeneous across the clusters (Izenman, 2008).

LCA is applied with built environment data as well as the travel behaviour data, as discussed in more detail in subsection 3.2.2 and subsection 3.2.1. Thus, two separate LCA's will be conducted to answer the RQ.

Various methods for conducting a cluster analysis exist in the literature. Latent Class Cluster Analysis (LCCA) is selected as the preferred method, as it outperforms other clustering methods like the k-means clustering method (Magidson & Vermunt, 2002). For the LCCA we use the software package Latent Gold 6.0 developed by (Magidson & Vermunt, 2002). LCCA offers flexibility with nominal and continuous variables (Kaplan, 2004). Another advantage of the LCCA is its probabilistic approach, instead of the deterministic approach employed in k-means. This probabilistic method maximises heterogeneity across clusters (Lee et al., 2019). Furthermore, the data did not require standardisation when using Latent Gold(J. K. Vermunt & Magidson, 2016).

LCCA has several sub-variants. For the RQ's the focus is on examining the effect of external variables on cluster membership. To achieve this, the step-3 method is employed, as it is the most suitable LCCA approach for modelling the associations between classes and external variables or class indicators (J. K. Vermunt & Magidson, 2020). The step 3 method consists of 3 steps, elaborated in this section.

Step 1: Define classes
In this step, different classes are generated using the indicators described in subsection 3.2.1 and subsection 3.2.2. Using Latent Gold, between 1 and 10 classes are generated. For this step it is important to consider the optimal number of classes for further analysis. This can be achieved by considering several statistical indicators:

- Bayesian Information Criterion (BIC) is a model fit indicator which needs to be minimized (Lezhnina & Kismihók, 2022; Nylund-Gibson & Choi, 2018), however some literature also suggest the heuristic elbow method by selecting the number of clusters where the reduction of BIC seems to drop off (Flynt & Dean, 2016) The BIC is calculated with the formula below:

$$\begin{aligned} \text{BIC} &= k \cdot \ln(n) - 2 \cdot \ln(L) \\ \\ k &= \text{number of parameters} \\ n &= \text{number of respondents} \\ L &= \text{likelihood} \end{aligned} \tag{3.3}$$

- Log likelihood (LL) needs to be maximized (Morgan, 2015)
- p-value needs to be significant (Nylund-Gibson & Choi, 2018)
- The BVR (bivariate residual) is the parameter for interaction effects between indicators. If the BVR is below 3.84 no significant interaction effects occur in the dataset (Kroesen, 2019). However, in some travel behaviour datasets, interaction effects are unavoidable. Therefore the interaction effects could included in clustering for less classes or interpretability (Molin et al., 2016; Reboussin et al., 2008; Wang et al., 2022).
- As a rule of thumb, each class should have a minimum size of 5% of the sample (Nylund-Gibson & Choi, 2018).

The BIC is selected as the primary indicator for determining the optimal number of classes. However, in addition to the BIC, other indicators such as the BVR, p-value and minimum class size should be considered collectively. It is important to note that the LL is already included in the BIC computation, so therefore only the BIC is considered. Moreover, the interpretability of the results must be taken into account when selecting the optimal number of classes. Thus, the final decision on the number of classes is based on a reasoned evaluation of all selection indicators mentioned above.

Once the optimal number of classes is determined, a dataset is generated that includes the probability of each individual belonging to a class. The assignment of class membership is subsequently conducted in step 2 of the step-3 LCA.

Step 2: Assign individuals to classes
In this step, individuals are assigned to a class based on the probability of belonging to each class, as determined in step 1. The assignment can be performed using different assignment methods: modal, proportional or random assignment (J. Vermunt, 2010). Modal assignment involves assigning an individual to the class with the highest membership probability. Proportional assignment distributes membership probability more evenly. While Modal and random assignment provide distinct assignment of class membership. Proportional assignment distributes class membership smoother. (Bakk et al., 2013; J. Vermunt, 2010).

Regardless of the method chosen, the assigned class membership and true class membership will differ for certain individuals (Bakk et al., 2013; Weller et al., 2020). The fact that estimated class membership does not always align with the observed class membership must be taken into account in the conclusion of the study (Weller et al., 2020).

Furthermore, there are correction methods designed to improve the accuracy of classifying group membership. These methods include the BCH (Block, Croon en Hagenaars) method (Bolck et al., 2004) and the Maximum likelihood (ML) method (J. Vermunt, 2010).

The preferred correction method is ML because it is a recommended correction method for handling co-variables (Bakk & Vermunt, 2015). Additionally, ML is favoured because it leads to smaller errors compared to other correction methods. However, when continuous variables are included in the dataset, the BCH method must be applied as state-of-the-art method (J. Vermunt, 2010).

Step 3: Relationship between class membership and indicators
In this step, the class membership variable will be used to examine the relationship between built environment data and class membership. This analysis is conducted using multi-nominal logistic regression analysis (J. K. Vermunt & Magidson, 2016), which will identify predictors of class membership. External covariates are included to examine the cluster membership probabilities.

Class membership estimation and the inclusion of external variables are performed across several variables. First the relationship between BE cluster membership and TB cluster membership is examined. Next the relationship between TB covariates and TB clusters are examined, followed by evaluating BE indicators in relation to TB clusters. Subsequently the relationship between BE covariates and TB clusters is examined. Finally, the relationship among all variables and TB clusters combined are analysed. All mentioned steps are illustrated in Figure 3.3. Using these values from the estimates enables the possibility of calculating the probability of belonging to a cluster.

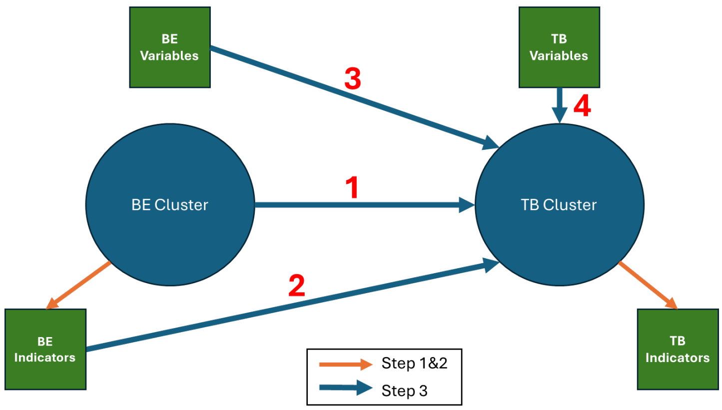


Figure 3.3: Visualisation of the analysis steps in the Step-3 LCCA. Red numbers represent the steps of the step-3 method. Blue lines are calculated in the third step and the orange lines in the first en second step of the step-3 LCCA.

4

Results

This section includes the execution of two LCCA's as outlined in chapter 3. Cluster models were developed with a specified number of classes. After selecting the most suitable number of cluster in the cluster model, these models, along with various external variables, were compared analytically and explained in spatial terms.

4.1. Travel behaviour LCCA

As outlined in chapter 3, the LCCA was conducted using travel behaviour data. Various indicators described in the data section were employed, but the LCCA solely on trips made, yielded the best statistical outcome. These results include the highest decrease in BIC, identification of logical classes and low BVR values (less interaction effects) and interpretability for further analysis.

Clusters ranging from 1 to 9 were generated, with statistics presented in Table 4.1. This table demonstrates a decrease in BIC and Log-Likelihood as the number of classes increases, with the reduction stabilising after 7 clusters, meaning after 7 clusters there is no increase in model fit. Table 4.1 also demonstrate that the BVR indicator, which captures interaction effects between clusters, reaches its lowest point in the 7-cluster model. Although the smallest cluster size contains 3.3% of the sample, the 7-cluster solution for the LCCA was selected due to its greater variation in public transport usage. Instead of having one cluster with public transport in the 5-cluster solution the 7-cluster solution contains public transport only, a combination of public transport and active modes and a mix of all modes. This offers more possibilities for the analysis of travel behaviour.

While some interaction effects are present in the 7-cluster solution, the maximum BVR in this solution is 13, which is the minimum from 1 to 9 cluster solutions. The n+1 and n-1 cluster models were not selected because the 6-cluster model has a high maximum BVR of 32, while the 8-cluster model contains a very small class representing 0.6% of the sample in addition the 8-cluster model also has a higher maximum BVR. In both the n+1 and n-1 clusters the BIC remains relatively stable.

	LL	BIC(LL)	Npar	Max. BVR	Class.Err.	Entropy R^2	Smallest class size
1-Cluster	-139135	278311	4	7970.5	0	1.000	1.000
2-Cluster	-123548	247189.2	9	10998.3	0.018	0.904	0.469
3-Cluster	-111080	222304.2	14	3349.7	0.027	0.919	0.166
4-Cluster	-106747	213690.6	19	27.4	0.041	0.893	0.122
5-Cluster	-103992	208231.7	24	24.3	0.066	0.859	0.121
6-Cluster	-103504	207307	29	32.3	0.076	0.839	0.034
7-Cluster	**-103226**	**206803.5**	**34**	**13.0**	**0.086**	**0.829**	**0.033**
8-Cluster	-103143	206689.1	39	16.3	0.093	0.820	0.006
9-Cluster	-103062	206577.1	44	16.5	0.096	0.817	0.004

Table 4.1: Travel behaviour latent class statistics with different number of clusters.

For each cluster the characteristics of the TB (travel behaviour) clusters is outlined below. This is succeeded by the examination of the inactive indicators, and the inactive TB covariates. The indicators are visible in Table 4.2, the covariates of travel behaviour in Table 4.3 and sociodemograpics in Table 4.4. Geographical distribution of the TB cluster membership of individuals is explained later in this chapter. Finally 7 clusters could be identified based on the trip indicators:

Cluster 1: Active only
25.6% of the sample, the largest cluster. The cluster is characterised by the use of active trips only. With 3.09 trips this is above the overall sample average. However, the total distance is way lower with 3.17 km per trip, likely due to the lower speed of active modes. Peak-hour trips are slightly above average and of course no car trips below 2 km were made by the respondents. The socio-demographics of this cluster show a lower mean age compared to overall sample, an average income level, slightly below average car ownership per household and a household size of 2.92, which is slightly above the mean.

Cluster 2: Car only
23.5% of the sample and is characterized by the car as only mode of transport with a average of 2.86 trips within a 24-hour period. The total distance travelled is higher than average with 56.2 km per day and 19.6 km per trip. 1/3 of the trips in this cluster are taken in peak-hours and 10% of the trips are under 2 km. The average age in this cluster is higher compared to the mean age, even as the income class. Car ownership is also higher than the average and the household size of 2.77 matches the overall average.

Cluster 3: Car+Active
20.9% of the sample and is identified by the use of both car and active modes of transport. With an average of 4.31 trip this is the highest trip count among all clusters. However, the total distance for both car and active trips is lower. The average car trip distance is 15,66 km, while for active trips the length is 2.73 km. This cluster has a high number of peak-hour trips. Although there are 0.5 less car trips, the number of car trips shorter than 2km is the same as the car only cluster.

Cluster 4: No trips
16.6% of the sample and is characterised by the fact that no trips are made within a 24 hour period. The mean age in this cluster is 52.3, which is 9 years older compared to the overall mean. This group also has the lowest income level among all clusters. Car ownership is below average, and the household size is the smallest of all clusters. This may be related to the older average age, as children have already left the household.

Cluster 5: Public transport + active
6.2% of the population and shows a mix of public transport trips combined with active trips. Notably, this does not include access and egress modes, as trips are aggregated by their main transport mode. The average number of trips in this cluster is 3,2 with a total distance of 45,2 km. Public transport trips have a distance of 26.8 km per trip, while active modes have a average of 2,73 km. There is a relatively

high number of peak-hour trips and as expected, no car trips under 2km are made. The average age is younger than the overall mean, the income level is below average and car ownership is also below average, with 0.8 cars per household.

Cluster 6: Public transport only
3,8% of the sample and is characterised by public transport trips only. The average distance traveled is 57.2 km, with an average trip length of 28.45 km, the highest trip length among all clusters. The number of peak-hour trips is around the overall mean. The mean age is similar with cluster 5 and also younger than the average age, the income class is exactly the same compared as in cluster 5.

Cluster 7: Mix PT, Car and Active
3,3% of the sample and features a mix of all modes. It has the highest total distance travelled, with an average trip length of 62.5 km. Although the number of trips is not the highest, the number of peak-hour trips are above average and car trips below 2km occurs at the average. The average age in this cluster is 38.6 years, younger than the overall mean. Income is the highest in the sample„ while car ownership and household size are both at the average.

	Active only	**Car only**	**car+ active**	**No trips**	**PT + Active**	**PT only**	**Mix PT, Car,Active**	**Overall**
Cluster Size	0.256	0.235	0.209	0.166	0.062	0.038	0.033	
Car trip	0	2.86	2.33	0	0.01	0	1.62	1.2122
PT trip	0	0	0.01	0	1.51	2.01	1.26	0.2138
Active trip	3.09	0	1.98	0	1.72	0	0.52	1.3294
made trip	1	1	1	0	1	1	1	0.8339
no trip	0	0	0	1	0	0	0	0.1661

Table 4.2: Indicators for cluster membership used in the LCA with their mean in each cluster.

	Active only	**car only**	**Car+ active**	**No trips**	**PT + Active**	**PT only**	**Mix PT, Car,Active**	**Overall**
Distance car	0	56.2	36.5	0	0.2	0	29.9	21.85
Distance PT	0	0	0.2	0	40.2	57.2	31.0	5.76
Distance active	9.8	0	5.5	0	4.7	0	1.6	4.01
total distance	9.8	56.2	42.2	0	45.2	57.3	62.5	31.62
peak-hour trips	1.12	1.02	1.42	0	1.20	0.96	1.19	0.98
Car 2km trip	0	0.30	0.29	0	0	0	0.14	0.14

Table 4.3: Travel behaviour variables per cluster with their mean in each cluster

	Active modes only	**Car only**	**Car+ active**	**No trips**	**PT + Active**	**PT only**	**Mix PT, Car,Active**	**Overall**
Age	39.9	46.6	43.2	52.3	36.7	36.4	38.6	43.84
Income class	6.6	7.2	7.2	5.7	6.2	6.2	6.8	6.69
HH car	1.1	1.6	1.5	1.0	0.8	0.8	1.2	1.25
HH size	2.92	2.77	2.93	2.51	2.63	2.65	2.77	2.79

Table 4.4: Travel behaviour socio-variables per cluster with their mean in each cluster.

4.2. Built environment LCCA

As outlined in section 3.2, the LCCA for the built environment was conducted using data from different sources. Several indicators described in the data section were used, the LCA that uses the proximity score, distance to train stations. Neighbourhood car ownership and job/household mix resulted in the best statistical outcome. This outcome results in the largest decrease in BIC, the identification of logical classes and low BVR values.

Clusters ranging from 1 to 9 were generated, resulting in statistics presented in Table 4.5. This table demonstrates a decrease in BIC and Log-Likelihood as the number of classes increases with the reduction in BIC stabilizing after 7 clusters. The BVR indicator, which captures interaction effects between clusters, reaches a low point in the 7-cluster model. However some interaction effects remain significant. The smallest cluster size contains 9.1% of the sample, which is sufficient for interpretability.

Although some interaction effects remain in the 7-cluster solution, the maximum BVR is 13.6, with 2 significant interaction effects. The n+1 and n-1 cluster models were not selected due to interpretability issues in the 6-cluster model and a small class size of 4% in the 8=cluster model. In both the n+1 and n-1 models, the BIC remains relatively stable. The results of the 7-clusters, further elaborated in this chapter, also shows some interpretable outcomes.

	LL	BIC(LL)	Npar	Max. BVR	Class.Err.	Entropy R^2	Smallest class size
1-Cluster	-5009	10069	8	369.9	0	1	1
2-Cluster	-4425	8957	17	152.1	0.065	0.797	0.481
3-Cluster	-4247	8655	26	56.5	0.072	0.834	0.238
4-Cluster	-4147	8513	35	48.2	0.124	0.787	0.177
5-Cluster	-4090	8454	44	29.4	0.145	0.775	0.121
6-Cluster	-4038	8406	53	13.8	0.141	0.795	0.0998
7-Cluster	**-3993**	**8373**	**62**	**13.6**	**0.144**	**0.803**	**0.091**
8-Cluster	-3964	8370	71	12	0.115	0.834	0.039
9-Cluster	-3934	8367	80	7.3	0.118	0.841	0.053

Table 4.5: Latent class statistics with a variable number of clusters. The 7-cluster model is the chosen model containing 7 clusters in total.

In Table 4.6, all indicators used to form the 7-cluster model are displayed alongside their mean values within each cluster. Each cluster has his own characteristics, described later in this section. The "overall" column represents the average across all clusters, this is the average over the 512 clusters used in the analysis. It must be mentioned that this is not the respondents data but the data of 512 postcodes.

	Cluster1	**Cluster2**	**Cluster3**	**Cluster4**	**Cluster5**	**Cluster6**	**Cluster7**	**Overall**
	Suburban train	Midsized cities TOD urbanization	Greenports and distant residential area	Distant cultural landscape	Logistics and ribbon development	Large cities	Residential urban sprawl	
Cluster Size (% of all postcodes)	0.254	0.150	0.147	0.133	0.130	0.095	0.091	
Score	103.85	158.53	89.35	49.93	43.34	252.46	162.40	114.23
Train station distance	3.07	1.62	9.54	19.85	7.38	1.80	4.36	6.60
Car ownership	1.00	0.66	0.99	1.19	1.20	0.44	0.66	0.92
Function-mix in range	1.18	1.29	0.94	0.96	1.54	1.48	1.00	1.19

Table 4.6: Indicators for cluster membership with their mean in each cluster

In Table 4.7, several socio-demographic variables of the built environments are presented, these are sociodemograpics from the respondents of the ODiN dataset, categorized by cluster. The spatial distribution of the clusters is shown in Figure 4.1. The names of all clusters are based on their characteristics and best matching built environment main group.

BE cluster	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7	Overall
	Suburban train	Midsized cities TOD urbanization	Greenports and distant residential area	Distant cultural landscape	Logistics and ribbon development	Large cities	Residential urban sprawl	
Age	44.99	42.19	45.25	47.16	44.61	37.42	45.2	43.84
Income class	7.14	6.06	6.99	7.14	7.34	5.65	6.13	6.69
Cars in HH	1.41	0.92	1.45	1.61	1.62	0.7	1.01	1.25
Size HH	2.94	2.46	2.95	2.89	3.08	2.41	2.63	2.79
ODiN Respondents percentage	34.8%	18.4%	14.5%	5.7%	7.5%	10.4%	8.7%	
Actual population percentage	32.2%	18.3%	14.7%	5.9%	7.1%	12.5%	9.2%	

Table 4.7: Sociodemographics per built environment cluster derived from the ODiN respondents. The ”respondents percentage” column indicates the percentage of the total ODiN respondents, while the ’populations percentage’ indicates the percentage of the total population in Zuid-Holland. So 32.2% in cluster one means that 32.2% of the population lives in cluster 1.

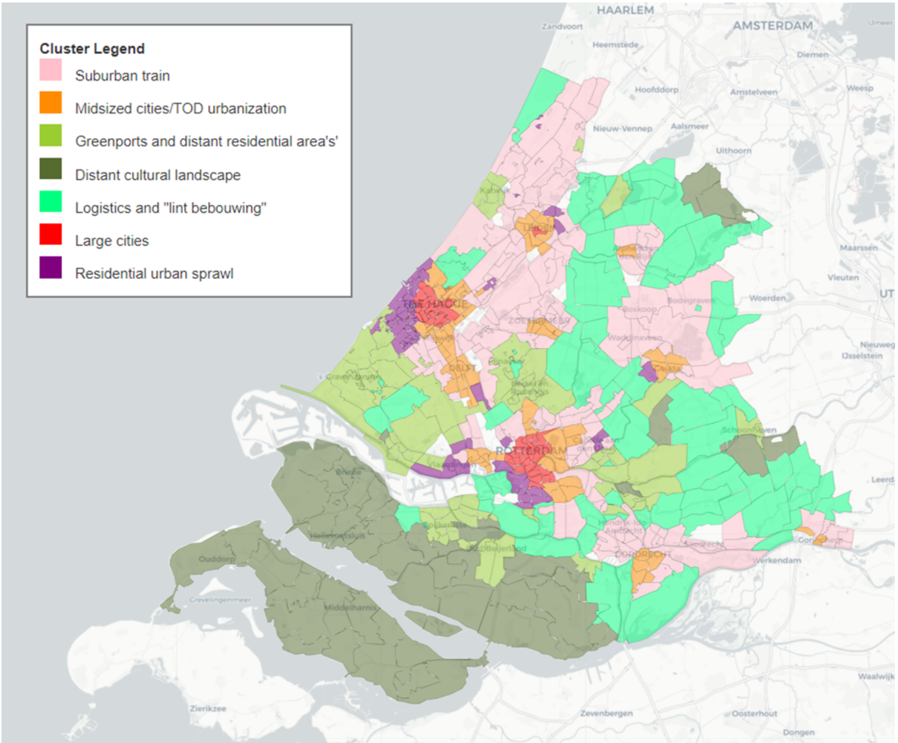


Figure 4.1: Built environment clusters

Cluster 1: Suburban 1 | Suburban train

This cluster compromises 25.4% of the postcodes, 32% of the population and 35% of the ODiN respondents. It is characterised by its proximity score around the median of the dataset, with a average distance of 3 km to train stations, making it accessible by bicycle. Car ownership is relatively high and the job/household balance is average. The average income class in this cluster is above average even as the car ownership per household. Spatially this cluster is mainly located in between the big cities and adjacent to cluster 2 and 7. Areas in this cluster could not be characterized as a city or rural area, so this is the first cluster characterized as suburban cluster.

Cluster 2: Suburban 2/City 1 | Midsized cities/TOD urbanization

Cluster 2 contains 15% of all postcodes and accounts for 18.3% of the population and 18.4% of the ODiN respondents. This cluster is defined by its short distance to train stations, higher proximity score, below-average car ownership and a higher number of jobs within bicycle distance. Spatially this cluster is mostly located near the borders of big cities (Rottedam, Den Haag and Leiden) or in mid-sized cities like Dordrecht, Delft, Gouda, Gorinchem and Alphen aan de Rijn. This cluster has a slightly younger population, lower income class, lower car ownership and smaller household size compared to the overall average. Some parts of the areas in this cluster are a midsized city while other parts are adjacent to the city. Therefore this cluster could be characterized as a partly city and suburban cluster.

Cluster 3: Suburban 3/Rural 1 | Greenports and distant residential area
Cluster 3 covers 14.7% of the postcodes, 14.7%population and 14.5% of the ODiN respondents. It is characterized by a below-average proximity score, greater distance to train stations (light-rail and metro not included see chapter 7), higher car ownership and a lower job/household ratio. Spatially, this cluster is located at a distance but not to distant form urban cores. Unless the metro is not included in the train station distance,it is notable that cities that are connected with light-rail or metro connections are partly assigned to this cluster, examples are Spijkenisse, Hoek van Holland, Maassluis and Pijnacker. Sociodemograpics of this cluster show a higher average age with a above-average income class, more cars per household and larger households. Some areas of this cluster could be characterized as suburbans in relationship with a big city like Spijkenisse and Pijnacker while other areas reveal more rural characteristics like the Westland areas, therefore this cluster is categorised as a suburban and rural mix cluster.

Cluster 4: Rural 2 | Distant cultural landscape
Cluster 4 includes 13.3% of the postcodes, 5.9% of the population and 5.7% of the ODiN respondents. This is a difference, but it is due to the fact that this cluster contains rural areas, where there are fewer inhabitants per postcode. This cluster has a relatively low average proximity score, further away from train stations, has a higher car ownership and a below average job/household mix. Spatially, this cluster is categorised as rural. It is mainly located at a distance from urban cores and satellite imagery shows these areas are primarily agricultural landscape. The population in this area is older than average, with higher income levels. Car ownership is high averaging 1.61 cars per household.

Cluster 5: Rural 3 | Logistics and ribbon development
Cluster 5 includes 13% of the postcodes, 7.1% of the population and 7.5% of the ODiN respondents. It is characterised by a lower proximity score, similar distance to train stations as cluster 4 and a high job/household mix, which is notable. Spatially, this cluster includes large logistic centres in the Westland and Bleiswijk as well as agricultural landscapes with ribbon developments along dikes. The average age is comparable to the overall mean, while the income class is higher. Car ownership is similar to cluster 4. Household size is significantly larger compared to the overall average. This cluster is categorised as rural due to its distance to rural areas and lower proximity score to facilities.

Cluster 6: City 2 | Large cities
9,5% of the postcodes is represented in cluster 6, with 12,5% of the population and 10,4% of the respondents. It has the highest proximity score combined with a relatively close distance to train stations, the lowest car ownership and a high job/household ratio. Spatially, this cluster is located in the three largest cities of Zuid-Holland: Rotterdam, The Hague and Leiden. The population in this cluster is younger on average, with the lowest income and car ownership across all clusters. Household size is also smaller.

Cluster 7: Suburban 4 | Residential urban sprawl
The 7th cluster accounts for 9.1% of the postcodes with 9.2% of the population and 8.7% of the respondents. It has a proximity score comparable to cluster 2, but the distance to the train stations is greater. Also car ownership is similar to cluster 2 but lower than the overall mean. The function mix is lower compared to cluster 2. Spatially, this cluster is mainly located adjacent to cluster 6, but specifically in areas without train stations. The average age in this cluster aligns with the overall average, while income class is slightly below average. The household size is 2.63, which is marginally smaller than the overall mean. This cluster is mostly situated on the edges of large cities, therefore it is categorised as a suburban cluster.

4.3. Estimation of cluster membership

The probability to belong to a specific travel behaviour cluster is calculated in this section. It is done with different types of variables in the end of this chapter all variables are included in one model to estimate the probability of travel behaviour cluster membership. For every type of variables dummy persons with different variables and probability to classes were created. This allows to show the effects of the relationships.

4.3.1. TB cluster - BE cluster

In this section, the membership of the travel behaviour clusters is reflected on residential location of individuals. Table 4.8 shows the distribution of people living in a specific built environment class across all travel behaviour clusters. Each column represents 100% of the travel behaviour categories for a particular BE cluster. Table 4.8 reveals some interesting insights:

- In the ”suburban train” cluster, 30% of individuals belong to the ”car-only” cluster, which is relatively high compared to the average of 23%. In the other suburban clusters the car use is lower.
- Despite availability of train stations, the suburban train cluster has a low percentage of PT class memberships.
- In all BE clusters, at least 20% of the individuals belong to the ”only active mode” cluster. The midsized cities, Large cities and residential urban sprawl clusters have the highest share of ”only active mode” individuals. However these modes are not only dominant in dense urban area’s even in the ”distant cultural landscape” cluster, 20% of individuals are in this active mode group.
- The only car use cluster accounts for 38% and 34% of the respondents in ”Distant Cultural landscape” and ”logistics and ribbon development” clusters. Which illustrates that individuals in these areas rely more on cars compared to large cities where only 14% of the respondents are in the ”car only” cluster.
- The percentage of people in the no trips made cluster is higher in ”distant cultural landscape” and the ”residential urban sprawl” cluster, however this percentage is slightly higher compared to the average of 16.6%.
- The public transport cluster is almost absent in the ”distant cultural landscape” and ”logistics and ribbon development” cluster, while in large cities PT is used the most frequently compared to other clusters.

	Suburban train	Midsized cities /TOD urbanisation	Greenports and distant residential area’s	Distant cultural landscape	Logistics and ”ribbon development”	Large cities	Residential Urban sprawl
	Suburban 1	Suburban 2/City 1	Suburban 3/Rural 1	Rural 2	Rural 3	City 2	Suburban 4
Only active modes	27% (2834)	34% (1837)	27% (1145)	21% (350)	24% (535)	34% (1055)	33% (846)
Only car	30% (3149)	20% (1085)	30% (1315)	38% (640)	34% (755)	14% (438)	22% (559)
Car + Active	18% (1901)	14% (756)	19% (828)	18% (306)	19% (418)	10% (312)	13% (348)
No trip	16% (1617)	17% (909)	16% (695)	19% (316)	17% (374)	17% (541)	19% (479)
PT + Active	3% (261)	6% (352)	3% (110)	1% (25)	1% (30)	11% (330)	4% (116)
PT	4% (389)	7% (393)	3% (147)	2% (38)	3% (65)	11% (335)	7% (188)
Mix PT, Active and car	2% (199)	2% (131)	2% (73)	1% (19)	2% (38)	3% (88)	2% (51)

Table 4.8: Travel behaviour cluster membership given the built environment cluster, this is based on where the ODiN respondents live. (So each column is 100%)

Table 4.9 presents the class membership probability estimates. The intercept represents the base class membership probability, with a intercept of 0.90 for ’active only’ resulting in a logit probability of 0.259. This reflects the size of the ’active only’ cluster. The utility contributions in the table show the strength of the relationships between BE and TB clusters. For instance, living in the ”suburban train” clusters increases the utility of belonging to a car-related cluster with a utility contribution of 0.21. The probability of belonging to a PT cluster is reduced with a utility contribution of -0.24. In cluster 6, the probability of belonging to a PT cluster is much higher, with a utility contribution of 1.008. Using this table the initial posterior probability of belonging to certain classes could be calculated.

From Table 4.9 several insight are revealed:

- Despite the relatively short distance to train stations in ”suburban train/ suburban 1” clusters, the probability of belonging to a car-related cluster remains relatively high. 48% Of all people belong to a car-related cluster, compared to 24% in cities (city 2) and 56% in rural areas (rural 2). Additional analysis could be conducted to reveal more insights in the suburban train cluster car trips.

- The cluster suburban 2/city 2 and suburban 4 show a lower probability to belong to a car cluster. This are mainly locations adjacent to city clusters.
- Differences between car related TB classes and BE clusters exists. In rural areas the utility contribution for car related clusters is high while the PT cluster contributions is lower. In contrast, urban classes demonstrate a high PT utility contribution while car related clusters are lower.
- Rural areas demonstrate some probability of individuals belonging to an active mode cluster. However the probability is higher in urban clusters.
- The probability of no trips appears to be higher in rural areas cluster 4 ('rural 2') and cluster 7 ('suburban 4').

	TB CLuster	active only	car only	car+active	no trips	PT + Active	PT only	Mix PT,Car,Active
BEcluster	Intercept	0.9093	0.8039	0.686	0.5364	-0.7188	-1.0497	-1.1671
Suburban 1		0.0172	0.2185	0.2016	-0.0557	-0.2403	-0.176	0.0346
Suburban 2/City 1		0.0332	-0.4351	-0.2865	-0.1829	0.5607	0.2196	0.0909
suburban 3/ Rural 1		0.0065	0.2417	0.272	0.0013	-0.1993	-0.2665	-0.0557
Rural 2		-0.0799	0.6746	0.4	0.3299	-0.5655	-0.4728	-0.2864
Rural 3		0.0045	0.4621	0.3504	0.1409	-0.7755	-0.2497	0.0673
City 2		-0.0379	-0.8676	-0.6672	-0.1925	1.0008	0.5635	0.2009
Suburban 4		0.0563	-0.2942	-0.2703	-0.0411	0.2191	0.3817	-0.0516

Table 4.9: Model for class membership probability based on the step-3 model.

4.3.2. TB cluster - TB covariates

Table 4.10 shows the relationship between travel behaviour clusters and TB covariates. The data indicate that some of the covariates have positive effects while others have negative effects on cluster membership. Notably, the number of cars in a household significantly impacts the probability of cluster membership. This is particularly observable in the car-related and public transport-related clusters. For instance, households with cars have a high probability of being in a car cluster and a low probability of belonging to a public transport cluster. Age and income also play a role in determining the cluster membership. Income, in particular, has a strong influence in the "no trips" cluster.

To illustrate the influence of certain variables, 5 dummy persons were been created. Table 4.11 presents these individuals and their respective BE characteristics. Table 4.12 displays the probabilities of belonging to a particular class. These probabilities are calculated using the coefficients from Table 4.10. It can be observed that not owning a car makes the probability of belonging to a car cluster lower, while the probability of using public transport increases. Age and income also play a role in the probability of membership. For example, person 2 aged 80, without a car, has the same probability of belonging to a public transport cluster as person 4 who owns a car. However, person 2 also has a much higher probability of belonging to the "no trips" cluster, which is consistent with the trends observed in the TB data section.

Intercept	**active only**	**car only**	**car +active**	**no trips**	**PT + Active**	**PT only**	**Mix PT, Car,Active**	Wald	p-value
	0.8827	-0.5513	-0.5258	-0.2928	0.9192	0.0545	-0.4866	529.3012	4.10E-111
Covariates	**active only**	**car only**	**car +active**	**no trips**	**PT + Active**	**PT only**	**Mix PT, Car,Active**	Wald	p-value
Age	-0.0007	0.0133	0.0071	0.0233	-0.0165	-0.0115	-0.015	1201.019	2.90E-256
HHsize	0.1301	-0.0847	0.0185	0.167	-0.0982	0.0229	-0.1556	290.5003	8.90E-60
HHAuto	-0.1639	0.6903	0.5169	0.0722	-0.6921	-0.6642	0.2409	1486.531	4,4e-318
Income	-0.0075	0.0265	0.0424	-0.1229	0.0287	0.0076	0.0251	403.9415	4.00E-84

Table 4.10: Model for TB clusters and Travel behaviour covariates of individuals.

	Age	Household Size	Household car(s)	Income class
Person 1	35	4	2	7
Person 2	80	2	0	5
Person 3	20	1	0	5
Person 4	50	2	1	9
Person 5	40	1	0	8

Table 4.11: Characteristics of dummy persons with travel behaviour covariates.

	active only	car only	car +active	no trips	PT + Active	PT only	Mix PT, Car,Active	prob. PT cluster
Person 1	24%	27%	27%	14%	3%	2%	3%	8%
Person 2	27%	15%	13%	34%	6%	4%	1%	12%
Person 3	32%	10%	11%	10%	21%	10%	5%	37%
Person 4	26%	25%	23%	13%	6%	4%	4%	13%
Person 5	31%	14%	14%	10%	18%	9%	4%	31%

Table 4.12: Probability of cluster membership for different dummy persons.

4.3.3. TB cluster - BE indicators

Table 4.13 shows the relationship between TB membership and the BE indicators used for the TB LCA. The function mix indicator (job/household ratio) was found to be insignificant in estimation of cluster membership and is therefore excluded from further analysis. In the estimation of TB cluster based on BE indicators it becomes clear that neighbourhood car ownership plays an important role in determining the probability of cluster membership. Although other relationships are weaker, they still indicate a consistent direction. For instance, a higher proximity score increases the probability of belonging to an active or public transport cluster, while a lower proximity score increases the probability of belonging to a car cluster.

These relationships are also demonstrated by using some sample persons. All variables differ between the sample persons. There are clear differences between sample persons and the different clusters membership probabilities. Car ownership, in particular, has a substantial influence on the probability of cluster membership. Although the proximity score is also important in determining the cluster membership probabilities. For example, a higher proximity score means a higher probability of being in the PT-only cluster and a lower probability of being in the car-only cluster.

To illustrate, consider person 5, who currently has a total probability of belonging to a public transport cluster of 9%. If his score increases to 250, his probability of being within a public transport cluster increases to 12.5%. If at the same time the number of cars in the neighbourhood is reduced to 0.8, the probability of belonging to a public transport cluster further increases to 17%.

Intercept	active only	car only	car +active	no trips	PT + Active	PT only	Mix PT, Car,Active	Wald	p-value
	0.5557	0.1719	-0.1076	0.8076	0.6337	-0.9938	-1.0674	51.3706	2.50E-09
Covariates	**active only**	**car only**	**car +active**	**no trips**	**PT + Active**	**PT only**	**Mix PT, Car,Active**	Wald	p-value
carown	0.2543	1.0921	1.0843	-0.1818	-1.7827	-0.4164	-0.0498	237.7364	1.70E-48
treinstationkm	-0.0014	0.0075	0.0043	0.0161	-0.0032	-0.0087	-0.0145	38.0369	1.10E-06
score	0.0012	-0.0029	-0.0012	-0.0019	0.0016	0.0027	0.0004	77.9253	9.60E-15

Table 4.13: Model for TB clusters and BE indicators

	Carownership Neighbourhood	Trainstation	Proximity score
Person 1	1.5	3	100
Person 2	1.2	10	60
Person 3	0.4	1	250
Person 4	0.4	5	150
Person 5	1.1	1	100

Table 4.14: Characteristics of dummy persons with BE indicators.

	active only	car only	car +active	no trips	PT + Active	PT only	Mix PT, Car,Active	prob. PT cluster
Person 1	21%	34%	30%	11%	1%	2%	2%	5%
Person 2	20%	32%	26%	15%	2%	2%	2%	6%
Person 3	32%	11%	13%	16%	17%	7%	4%	29%
Person 4	27%	15%	14%	20%	14%	5%	4%	23%
Person 5	24%	28%	25%	14%	3%	3%	3%	9%

Table 4.15: Probability of cluster membership for different dummy persons using the BE indicators.

4.3.4. TB cluster - BE covariates

In this analysis, the travel behaviour clusters were compared with the BE covariates, all measured at postcode level. Surprisingly, it can be seen that address density does not play a major role in the probability of belonging to a particular cluster. However, the percentage of multi-family houses does have an impact on cluster membership. Areas with more multi-family houses are associated with less car use and higher use of public transport.

Distances to highways are short on average, making this value a relatively minor factor in determining cluster membership. However, the average household income in a neighbourhood appears to have a strong influence on the probability of belonging to a particular cluster. A higher neighbourhood income increases the probability of belonging to a car-related cluster and decreases the probability of belonging to a public transport cluster, and vice versa.

These relationships are similar to the influence of individual income class discussed earlier. House value (WOZ waarde) also plays a role in class membership. The higher the value of the house, the higher the probability of being a member of a public transport cluster and the lower the probability of being a member of a car cluster.

Intercept	active only	car only	car +active	no trips	PT + Active	PT only	Mix PT,Car,Active	Wald	p-value
	0.9575	1.1647	0.9685	1.0377	-0.7566	-1.9202	-1.4517	228.9575	1.30E-46
Covariates	active only	car only	car +active	no trips	PT + Active	PT only	Mix PT, Car,Active	Wald	p-value
Adress density	0.0001	-0.0001	-0.0001	-0.0001	0.0001	0	0	110.8738	1.30E-21
Percentage multi-family homes	-0.0036	-0.0076	-0.0081	-0.0015	0.0093	0.008	0.0035	142.5099	3.00E-28
Highwayramp km	-0.0004	0.012	-0.0124	0.0323	-0.0664	0.0929	-0.0579	15.0534	0.02
Average household income	-0.0196	0.0376	0.013	0.011	-0.0643	0.0281	-0.0058	104.9452	2.30E-20
Woz worth	0.002	-0.003	-0.0004	-0.0023	0.0049	-0.0023	0.0012	211.4229	7.00E-43

Table 4.16: Model for TB clusters and BE covariates.

Sample individuals also show clear differences in travel behaviour. Especially in the car-related clusters and the active clusters, there are large differences between the sample persons. The other clusters show relatively small variations between the sample persons. It is notable that person 1 has a low probability of belonging to an PT cluster while living in a high density area.

	Address density	Percentage multi-family homes	Distance highway ramp	Woz worth (*1000)	Average household income
Person 1	2500	15	1	600	45
Person 2	1000	60	2	250	25
Person 3	6000	40	4	200	30
Person 4	3000	60	4	400	40
Person 5	2500	20	1	350	35

Table 4.17: Characteristics of dummy persons with BE covariates.

	active only	car only	car +active	no trips	PT + Active	PT only	Mix PT, Car,Active	prob. PT cluster
Person 1	39%	18%	23%	8%	6%	1%	3%	11%
Person 2	25%	24%	19%	20%	6%	3%	3%	12%
Person 3	36%	24%	14%	15%	4%	4%	2%	11%
Person 4	31%	23%	18%	15%	5%	5%	3%	13%
Person 5	29%	26%	22%	14%	4%	2%	3%	9%

Table 4.18: Probability of cluster membership for different dummy persons with BE covariates.

4.3.5. TB cluster - All indicators and variables combined

This comparison includes all indicators, both of the respondents and of their built environment. In this way the strength and direction of the various relationships become more visible. For example, it can be seen that the car has a strong influence on travel behaviour. Other social variables also influence the probability of cluster membership, as well as some built environment indicators.

In the profiles of the 5 individuals, the estimated class probabilities become even more clear. For example, the probability of belonging to a car-related cluster may be strongly reduced both by low car ownership in the area and by the individual (there may be some correlation), while the probability of belonging to a PT cluster increases under the same conditions. Additionally, we can observe that other factors also influences the probability of belonging to a particular cluster. Interestingly, these factors appear to play a smaller role in car clusters compared to public transport clusters where the difference between sample individuals is relatively small.

Intercept	active only	car only	car +active	no trips	PT + Active	PT only	Mix PT,Car,Active	Wald	p-value
	1.1323	-0.5282	-0.0336	-0.066	1.3379	-0.9296	-0.9127	60.4238	3.70E-11
Covariates	**active only**	**car only**	**car +active**	**no trips**	**PT + Active**	**PT only**	**Mix PT, Car,Active**	Wald	p-value
Respondents									
Age	-0.0006	0.0112	0.0057	0.0223	-0.0124	-0.0113	-0.0148	928.85	2.20E-197
Household size	0.1292	-0.1162	-0.0014	0.1514	-0.0355	0.0297	-0.1572	294.8332	1.00E-60
Income class	-0.0136	0.0275	0.0363	-0.1176	0.0303	0.0114	0.0258	349.8453	1.70E-72
Household car(s)	-0.1767	0.5948	0.4452	0.027	-0.5281	-0.6298	0.2676	1114.113	1.80E-237
BE									
Proximity score	0.001	-0.0016	-0.0004	-0.0004	0.0019	-0.0002	-0.0003	18.1085	0.006
Percentage multi-family homes	-0.0045	0.0004	-0.0048	-0.0003	-0.0026	0.0053	0.0066	23.8326	0.00056
Average household income	-0.016	0.0019	-0.0022	-0.0007	-0.0113	0.0408	-0.0125	13.2245	0.04
Highwayramp km	0.0028	-0.0211	-0.0397	0.0043	-0.0332	0.1269	-0.04	17.63	0.0072
WOZ worth	0.0019	-0.0012	0.0003	-0.0008	0.0016	-0.0031	0.0014	54.346	6.30E-10
Trainstation km	0.0021	0.0042	0.0058	0.0117	0.0028	-0.0161	-0.0105	16.6115	0.011
Carownership Neighbourhood	-0.1987	0.9168	0.0884	0.2149	-1.4232	0.1472	0.2546	37.8003	1.20E-06

Table 4.19: Model for TB clusters with all variables used.

	Person 1	Person 2	Person 3	Person 4	Person 5
Age	35	80	20	50	40
Household Size	4	2	1	2	1
Income class	7	5	5	9	8
Household car(s)	2	0	0	1	0
Proximity score	100	60	250	150	100
Percentage multi- home	15	60	40	60	20
Average household income	45	25	30	40	35
Highway-ramp km	1	2	4	4	1
House value (WOZ)	600	250	200	400	350
Train-station km	3	10	1	5	1
Carownership Neighbourhood	1.5	1.2	0.4	0.4	1.1

Table 4.20: Characteristics of dummy persons with all indicators included.

	active only	car only	car +active	no trips	PT + Active	PT only	Mix PT, Car,Active	prob. PT cluster
Person 1	27%	28%	29%	10%	1%	1%	4%	6%
Person 2	20%	23%	12%	38%	2%	3%	1%	7%
Person 3	32%	6%	9%	9%	26%	14%	3%	43%
Person 4	31%	16%	21%	12%	12%	5%	4%	20%
Person 5	32%	19%	18%	11%	9%	7%	4%	20%

Table 4.21: Probability of class membership for different dummy persons using all indicators.

.

4.4. Analysis of travel behaviour

In this section the differences in travel behaviour across the BE clusters are analysed in more detail. Travel behaviour is studied based on the distance of work trips Table 4.22 and car share of commuter trips Table 4.23 and Table 4.24. This type of trips is chosen because it is one of the most frequent purposes in number (18% of all trips). The trips are also taken at an interesting time, often in the morning or evening peak hours. Data presented in Table 4.22, Table 4.24 and Table 4.23 reveals some insights. From these tables it could be seen that:

- Until 2 km there isn't much difference in the percentage of work-trips taken, this distance is ideal for walking (Table 4.22).
- In the cities cluster 30-50% more people live within a 5 km distance from their work location. In cities the percentage is 36% and in the other clusters around 25%. This distance could be ideal for cycling (Table 4.22).
- The distance from 5-15 km contains 30% of all suburban trips (cluster 1) and 21% of all large city trips, this distance is ideal for the e-bike (Table 4.23).
- From the trips until 5km the work related car trips under 5km are +-30% in the suburban cluster 1 and rural cluster 4. While the car share is 9% in the 'large cities' cluster (Table 4.23).
- Circa 60% of the work related trips in cluster 1 'Suburban Train' and cluster 6 are within 15 km (Table 4.22), so the majority of work trips occur within a relatively short distance. Even rural areas, cluster 4, has 45% of the work-trips within 15km. This trips are ideal to be made with an E-bike.
- Car use shows some variation across clusters. In cluster 1 'Suburban Train' 43% of all work-trips below 15 km are made by car, while in cluster 6, the 'large city' cluster only 21% of the trips below 15 km are made by carTable 4.23. The most rural area is higher, with 50% of all work-trips below 15 km made by car.
- Zooming in to the model split between distances of 5-15 km (Table 4.24) reveals that the share of car is 34-62.5% in suburban area's while the model split is 62-75% in rural and 35% in urban areas.
- Table 4.24 shows some variation in car use between suburban clusters however the car use lies between large cities and rural areas.
- Active modes of transportation are occur in all clusters. Large cities and residential urban sprawl clusters show the highest active mode share. However, also rural an suburban BE clusters show commuter trips taken by active modes. However the share is lower in suburban cluster with 81.5% of work-related trips until 2 km by active modes compared to 93% of all active mode trips in large cities section E.4.

Trips	Cluster 1: Suburban Train	Cluster 4: Distant cultural landscape	Cluster 6: Large cities
until 2km	11% (n = 574)	15% (n = 132)	14% (n = 271)
until 5km	28% (n = 1471)	25% (n = 226)	36% (n = 690)
until 15 km	58% (n = 3113)	45% (n = 409)	58% (n = 1103)
Total	100% (n = 5365)	100% (n = 908)	100% (n = 1907)

Table 4.22: Distance of work related trips in 3 clusters: suburban train. distant cultural landscape and large cities. Note: This are trip distances, no radius. (CBS, 2019, 2020b, 2024a). n represents the number of trips.

Car Trips	Cluster 1: Suburban Train Car (%)	Cluster 4: Distant cultural landscape Car (%)	Cluster 6: Large cities Car (%)
until 2km	18 (n = 105)	16 (n = 21)	5 (n = 14)
until 5km	29 (n = 429)	31 (n = 70)	9 (n = 60)
until 15 km	43 (n = 1336)	51 (n = 207)	18 (n = 203)
Total	58 (n = 2977)	70 (n = 630)	29 (n =465)

Table 4.23: Car related work trips in each area. The percentage is cumulative and taken as percentage of all trips in the trip kilometer class. (n = xx) represents the number of car trips.

Commuter trips with car	**Suburban Cluster 1**	city /Suburban Cluster 2	rural /Suburban Cluster 3	**rural Cluster 4**	rural Cluster 5	**Large cities Cluster 6**	city /suburban Cluster 7
0 - 2 km	18.3% (n = 105)	8.8% (n = 32)	15.9% (n = 52)	15.9% (n = 21)	19.1% (n = 26)	5.2% (n = 14)	13.4% (n = 19)
2 - 5 km	36.1% (n = 324)	18.9% (n = 119)	41.7% (n = 130)	52.1% (n = 49)	37.2% (n = 42)	11.0% (n = 46)	23.4% (n = 71)
5 - 15 km	55.2% (n = 907)	43.9% (n = 311)	62.5% (n = 423)	74.9% (n = 137)	62.4% (n = 239)	34.6% (n = 143)	34.0% (n = 150)
Total	55.8% (n = 2995)	38.3% (n = 1152)	58.5% (n = 1348)	69.5% (n = 631)	65.2% (n = 749)	24.4% (n = 466)	37.9% (n = 505)

Table 4.24: Car share within each cluster and distance category. The mode of these trips is commuting. (n = xx) represents the number of car trips.

4.4.1. Travel behaviour filtering on station distance in suburban areas

To further clarify the differences between suburban clusters and urban clusters, a filtering was conducted within the largest suburban (cluster 1), to areas with the same or a lower average distance to train stations compared to the urban clusters.

This filtering is shown in Table 4.26, where the modal split for cluster 1 is visible. Despite the proximity to stations, the modal split in filtered areas is almost comparable to the entire suburban cluster shown in Table 4.25. PT usage is 4 percentage points higher in the filtered areas however, this is not a result of a reduction in car usage but rather a consequence of the decrease in active modes of transportation.

For shorter distance the use of car is interesting in suburban areas with a proximity to train stations (and to other facilities). Do the residents use the car while there are alternatives to car use or do the have to use the car due to their inaccessible work location, time of work or simply a car minded attitude? For the scope of this research this isn't investigated. However the higher car use in suburban clusters, even for short distances and with possibilities for other modes of transport, is an interesting topic for further research.

	Car	**PT**	**Active**
2km	18% (n = 105)	0% (n = 1)	82% (n = 468)
5km	29% (n = 429)	2% (n = 32)	69% (n = 1010)
15km	43% (n = 1336)	8% (n = 237)	49% (n = 1536)
total	56% (n = 2977)	14% (n = 745)	30% (n = 1595)

Table 4.25: Suburban train cluster modal split (cumulative)

	Car	**PT**	**Active**
2km	26% (n = 45)	0% (n = 0)	74% (n = 125)
5km	33% (n = 142)	3% (n = 12)	64% (n = 272)
15km	44% (n = 389)	10% (n = 91)	46% (n = 402)
Total	55% (n = 859)	18% (n = 274)	27% (n = 418)

Table 4.26: Suburban train (cumulative): only postcodes within a distance of 1.8 km or less from train stations are included.

4.4.2. Car ownership and model split of the car in commuter trips

Car ownership appears to play a significant role in travel behaviour. This section examines the effect of household car ownership on the share of car use. Focusing on whether individuals in suburban clusters use their cars more frequently compared to individuals in urban clusters and how this varies with the number of cars owned. In this analysis commuter trips are analysed as they are among the most common trip purposes and typically occur during peak hours, making them interesting for understanding travel beahviour.

In Table 4.27 and Table 4.28, the share of car use is presented for each cluster and household car ownership. The share of car use is as given as percentage of the total trips (n). Table 4.27 shows car use for all commuter trips, while Table 4.28 focusses on commuter trips shorter than 5 km. As noted in chapter 3 the car category contains car, motor-vehicles and delivery vans. Trips made with the moter-vehicle and delivery van are excluded from the '0' car ownership column as they do not reflect actual car use

It is surprising that even individuals without a car in their household travel by car. More surprising is the difference in car use between clusters. For instance in suburban cluster 1, individuals with at least one car use their car for 48% of all trips, compared to 35% in urban cluster 6. For shorter trips (<5 km), this difference is even more visible: 24% in suburban cluster 1 versus 12% in urban cluster 6. This findings highlight the higher car use in suburban areas even when households own a car.

Table 4.29 provides additional insights by comparing the number of cars in a household to the number of trips. The data reveals a clear trend: As the number of cars in a household increases, the total number of trips, is increasing.

Car ownership	0	1	2	3+
Cluster 1	18% (n=404)	48% (n=2430)	68% (n=1939)	75% (n=564)
Cluster 2	8% (n=745)	39% (n=1618)	67% (n=516)	67% (n=126)
Cluster 3	14% (n=163)	54% (n=1008)	68% (n=812)	71% (n=312)
Cluster 4	23% (n=31)	61% (n=341)	76% (n=387)	79% (n=146)
Cluster 5	No Data (n=25)	53% (n=436)	71% (n=469)	82% (n=217)
Cluster 6	4% (n=848)	35% (n=843)	60% (n=179)	50% (n=36)
Cluster 7	7% (n=322)	39% (n=683)	63% (n=257)	60% (n=68)

Table 4.27: Car ownership and model split of the car (n represents total number of trips)

Car ownership	0	1	2	3+
Cluster 1	7% (n=151)	24% (n=730)	36% (n=447)	53% (n=143)
Cluster 2	3% (n=314)	18% (n=527)	32% (n=119)	no data (n=28)
Cluster 3	4% (n=49)	28% (n=329)	32% (n=191)	36% (n=70)
Cluster 4	no data (n=11)	25% (n=106)	35% (n=81)	no data (n=28)
Cluster 5	no data (n=12)	31% (n=123)	23% (n=78)	31% (n=36)
Cluster 6	2% (n=327)	12% (n=288)	21% (n=58)	no data (n=16)
Cluster 7	7% (n=123)	24% (n=224)	32% (n=77)	no data (n = 21)

Table 4.28: Car ownership and model split of the car for trips <5 km (n represents total number of trips <5km)

Car in HH	0	1	2	3+
Respondents	5524	14414	7756	2302
Trips per person	2.53	2.91	3.11	3.14
Commuter trips	0.46	0.51	0.59	0.64

Table 4.29: Car ownership and number of trips per person and for commuting

5
Conclusion

In recent years, car ownership in Zuid-Holland has increased and it is expected to continue increasing in the future. Since growth in car ownership varies regionally, it is interesting to distinguish different built environments and examine their influence on car-related travel behaviour. Therefore, this study explores travel behaviour across different built environments, with a specific focus on suburban areas, areas that can't be classified as rural or urban areas. The main question in this study is: Which mobility patterns can be identified and to what extent are they influenced by built environment?

Several clusters of the built environment can be identified in the province of Zuid-Holland. Seven clusters have been identified based on homogeneity within clusters and heterogeneity between clusters. Four clusters show suburban characteristics. These suburban clusters differ from each other on the basis of the indicators used. This is consistent with the theory of suburban areas (Forsyth, 2012; Hamers, 2003). However, their characteristics remain between those of urban and rural clusters. The remaining clusters show rural or urban characteristics, making them suitable for comparison with suburban areas.

The comparison of BE (built environment) cluster membership with the probability of belonging to a particular TB (travel behaviour) cluster shows that in suburban and rural areas the probability of belonging to a car cluster is higher, while the probability of belonging to a public transport cluster is lower. In urban areas the opposite is observed. It can be observed that active clusters are present everywhere, although the probability of belonging to an active cluster is higher in urban clusters.

Comparing suburban areas with other area types provides valuable insights into differences in travel behaviour. Data analysis shows that car use is much higher in suburban areas, even for short distances. The distance to stations does not play a visible role. The filtering of suburban areas at the same or lower distance from stations compared to urban clusters, results in similar travel behaviour compared to the unfiltered suburban area. Furthermore, the analysis of travel behaviour data and BE clusters shows that suburban areas are located between rural and urban areas. In the urban clusters travel behaviour is less car oriented and in the rural clusters travel behaviour is more car oriented. In terms of travel behaviour, the share of work trips made by car is 24% in urban clusters, 70% in rural clusters, and between 38% and 58% in the 4 suburban clusters. Travel behaviour differs across suburban clusters, where some areas are more similar to rural regions, while others tend toward urban patterns. It is interesting to see that despite relative good proximity to facilities like supermarkets, shops and train stations, suburban areas reveal a significantly higher car use compared to urban regions.

Even when someone owns a car, suburban areas show a higher car use compared to urban areas, even for short distances. In one suburban cluster, the share of work trips by car for households owning one car is 48% compared to 35% in urban clusters. For short trips under 5 km, this difference is 24% and 12%. More trip data could provide further insight into the relationship between car ownership and mode choice.

As shown in the literature review, the suburban area is an interesting area for traffic behaviour studies, at the same time this area is a difficult area to define. This study provides a distinction of different areas including three types of suburban areas.

This provides insight into the different mobility patterns between the clusters based on travel behaviour data. The literature, including case studies conducted in the Netherlands, shows that there is little understanding of precise travel behaviour within suburban areas. This case study improves this understanding. In addition, a methodology has been developed to compare travel behaviour between BE clusters. This makes the study applicable to other areas as well.

This research provides more knowledge on travel behaviour in certain areas. This leads to valuable insights for mobility policies. One insight is that in many suburban areas the car is used for short distances while in urban areas the use of the car for short distances is much lower. So theoretically, less car use for short distances is possible in suburban areas. This provides opportunities for policies specifically targeting on these areas. With the estimations of probabilities for cluster membership, this research provides tools to predict which travel behaviour is expected in certain built environments. This would help to guide spatial developments. It improves the ability to estimate what kind of travel behaviour can be expected in a specific area. However, additional research is needed for more reliable cluster membership estimates.

Future research could include, for example, the implementation of a larger and more precise sample, since currently certain groups are not fully represented. In addition, it is also recommended to analyse travel behaviour of multiple days in the dataset. Currently only data from one day is included. It is recommended to consider not only residential but also destination locations. Currently, all travel behaviour is linked to the respondent's residential location, including destination locations in the analysis could provide interesting results.

To summarise, the study demonstrates that built environment influences expected travel behaviour. In urban areas, there is a high probability of an PT-related mobility pattern, while in suburban and rural areas, the probability of car-related mobility pattern is higher. The data analysis reveals that suburban areas represent an intermediate category in terms of travel behaviour. Suburban areas are neither comparable to urban or rural areas. This aligns with general statements about suburban areas: they are neither rural nor urban (Forsyth, 2012; Hamers, 2003).

6 Discussion

The results of this study show that car use is significantly higher in suburban areas, even for short trip distances. Another remarkable result is that active modes of transport are present in all clusters. These findings are consistent with the research by (Liu & Alain, 2014; Vega & Reynolds-Feighan, 2009b; Wiersma, 2020), which also show that suburban areas have a car-oriented mobility pattern. It is remarkable that in terms of the proportion of distances for work trips up to 15 kilometres, the distances cover the same proportion in urban and suburban clusters. However, travel behaviour in these clusters is completely different.

A possible explanation is the greater distance to stations. However, even if the distance to stations is the same as the urban area, travel behaviour differs. A possible explanation is that in suburban areas stations are often served by lower frequency and slower sprinter trains, while in urban clusters intercity trains depart more frequently. This difference in accessibility may influence the choice of transport mode.

Individual car ownership plays an important role in the relationship between the built environment and the probability of TB cluster membership, as previous findings have shown (Ding et al., 2017; Van Acker & Witlox, 2009). The relationship with other built environment factors is weaker but present. In individual profiles, there is a distinction visible in the probability of cluster membership. The probability of an car-oriented cluster for the dummy persons is heavily weighted by individual car ownership.

A limitation of this study is that the CBS data did not include light rail, metro and tram lines in the calculation of distance to train stations. This may affect the classification of towns such as Spijkenisse, Hoek van Holland and Pijnacker. These areas are adjacent to light rail or metro lines , which are similar to sprinter stations, which are included in the station distance dataset. These sprinter stations are mainly present in the suburban areas. This may lead to local errors in classification.

Another limitation relates to the representation of TB data. It should be taken into account that the age, household size, income class and average car ownership in the sample is higher compared to the average in Zuid-Holland, logically this affects TB cluster membership. Furthermore, some groups are not fully represented such as non-Western migrants (CBS, 2024a). CBS uses stratification at national level to achieve a more correct weighted representation of person characteristics. In this study, stratification is not applied since individual trips were also considered. This offers an opportunity for future research (in combination with a more representative sample).

Furthermore, the travel behaviour of individuals was only measured over a 24-hour period, which means that the results at individual level could differ, as travel behaviour can vary from one day to the other. The study would be more robust if travel behaviour was analysed over several days. However, there's currently no data available for the province of Zuid-Holland that provides this information with a usable sample size.

An important point to consider is the influence of residential self-selection on the difference in travel behaviour between the BE clusters. Residential self-selection means that people choose their residential location based on their travel behaviour. For example, individuals who travel a lot by car are more likely to choose a residential location that is easily accessible by car. These individuals are less likely to live in a car-free city centre (van Wee, 2009).

Residential self-selection might contribute to the higher car use in suburban clusters. Those who prefer travelling by car are more likely to live in regions that support car usage, such as suburban or rural areas. Area types that typically implement fewer car-oriented policies, such as limiting parking spaces or introducing paid parking. However other factors like social, economical and geographical factors also play important a role in residential selection.

Besides residential self-selection, attitude may also play a role in higher car use in suburban areas. Residents of suburban areas may have a certain preference in the use of cars something which could be influenced by several factors (van Wee et al., 2019). The same attitude may conversely apply to public transport and active forms of transport in urban areas. In addition to residential self-selection, a car-minded attitude in suburban areas may be a factor behind their higher levels of car use (van Ham, 2012).

The influence of paid parking and parking policy on both destination and arrival location is not included in this study. Parking policies were not included due to the higher scale at Postcode 4 level. This scale level makes it difficult to perform precise analyses on parking policies because parking policies are often only implemented in parts of the postcode (www.prettigparkeren.nl, 2024). Therefore, parking policies were not included in this analysis. However, parking policies can play a role in transport mode choice (P. Christiansen et al., 2017).

Based on the points made earlier, several possibilities for further research expansion emerge. These are elaborated in the next section, the recommendations.

7 Recommendations

Based on the results of this study and (data) sources used for this purpose, some recommendations can be given for future research in addition to this thesis. The recommendations are divided into four categories: Data used, data processing, methodology improvements and policy recommendations.

7.1. Data used

Some improvements regarding the used ODiN data could be made. A significant improvement is that the sample size could be increased by using single-year data instead of combined data from different years. This increases the consistency of results. To achieve this, additional data can be collected throughout Zuid-Holland, as is already done in the MRDH (Metropoolregio Rotterdam-Den Haag) and Utrecht. This can be done in cooperation with CBS. In addition, increasing the sample size would also help to better represent under-represented groups. This is now solved by CBS by national stratification, but cannot be adopted 1:1 for Zuid-Holland. A more representative sample specifically for Zuid-Holland could further increase the reliability of the conclusions.

The scale level of ODiN data could be refined from the current Postcode-4 to Postcode-5 level. This would lead to more accurate measurements of distances to stations and facilities. Also parking policies could be included in the analysis at this smaller area level. After all, parking policy has a strong influence on mode choice (P. Christiansen et al., 2017). This might lead to more accurate estimates of cluster membership as the data for individuals in the ODiN data can be reported more precisely.

In CBS data by postcode area, the distance to stations were calculated as an average distance for all residents in the postcode. As mentioned earlier, this does not include light rail and metro. The metro network does fulfil a regional function in the MRDH. In this study, the distance was not included due to the method of calculation at personal level in combination with the desire to be able to make the study applicable to other areas in the Netherlands.

For future research, it is recommended to include light rail/metro connections and place them under the same indicator as distance to stations. This is partly because these modes of transport share a similar travel speed and travel on comparable regional route. Another reason is that travellers with these connections show almost similar travel behaviour. Metro/light-rail stations can therefore also be regarded as a high frequency suburban ”station” in the built-up area.

7.2. Data processing

In this study the parameters for calculating the proximity score are partly based on a survey of mobility policy experts within the province of Zuid-Holland. The values of marginal utility and facility scores, are currently based on this survey. These values might be investigated more extensively in the future, for instance through new/additional research to get a more accurate understanding of how proximity to different facilities is perceived.

Car ownership can also be measured in alternative ways. In addition to current car ownership per household, the number of cars per adult or per hectare can offer interesting insights. The number of cars per hectare has been applied in research by Bagley and Mokhtarian, 2002. These values offer possibility to better understand the effect of car ownership on both streetscape and travel behaviour.

Currently the mode of transport is only considered for trips with commuting as purpose. This analysis shows large differences in travel behaviour between BE clusters. The reason this motive was chosen is the fact that commuting trips are most frequent in number but also at an interesting time, often in the morning or evening peak hours.

Further research could investigate other travel motives like sports or shopping as well. This could lead to broader insights into travel behaviour and related built environment factors. This could be useful for policymakers. Due to the scope of this study, travel behaviour in suburban areas, a breakdown by motive was not performed, but the data has been prepared for further analysis.

7.3. Method improvements

In terms of methodology, the main focus was now on the respondent's residential location. The trips made by the respondent were summed and placed in the residential postcode of this respondent. Further analysis continued with the residential postal code during the study. Alternatively, this could be carried out looking at the trips' origin and destination postcodes. By basing a cluster analysis on the trips themselves and not just the individual, new insights into travel patterns can be obtained. In this case, the arrival and destination postcode of all trips is included as well. This provides opportunities for more in-depth analysis, for example, to travel behaviour along bottlenecks in the regional transport network. Clustering trips using the arrival and destination postcode is therefore a recommendation for future research.

In this study, a specific filtering was based on the distance to train stations. This was done to evaluate how effectively a provincial policy, building near stations in suburban areas performs in practice. For future research, additional policies could be tested using the extensive dataset on individual travel behaviour and socio-demographic characteristics. Examples include filtering by travel behaviour by variables such as car ownership, income group or age. Furthermore, the dataset contains additional respondent information that could be used to deepen the analysis.

7.4. Policy recommendations for the province of Zuid-Holland.

In order to ensure that the province of Zuid-Holland remains accessible to all residents, reducing car use is necessary, particularly in places where adequate alternatives to the car are available. In suburban areas, where facilities and public transport are relatively nearby, the car is used quite often, even for short distances. In this section, policy recommendations are provided to specifically reduce car use in these areas and promote more sustainable mobility options.

Guidance in developments near stations:
According to travel data from this research, car use is relatively high around station areas in suburban areas. Therefore, it is useful to set conditions for residential developments around station areas to achieve a low-car mobility patterns and improve the use of stations. This can be achieved with the following measures:

- Guidance on the type of housing to be built: This research shows that higher-income young adults are more likely to adopt an PT-related mobility pattern. Therefore, it is desirable to focus on developing housing for this target group around station areas. This means prioritizing more affordable starter homes instead of single-family houses. For rental housing, the focus could be on youth housing with conditions of a maximum rental period.
- Mandatory PT subscriptions for new housing near station areas: Similar to how residents are required to pay a waste collection fee, a mandatory PT subscription for new-home households near station areas could be considered in order to achieve less car oriented travel behaviour. Therefore it is recommended to further explore the implementation of this measure.

- Parking permits linked to a PT subscription: First parking permit zones have to be implemented near station zones. Subsequently, when issuing a parking permit in suburban station areas, a condition could be set that requires the permit holder to obtain a public transport subscription as ”fee” for the permit. This will ensure that every car in the neighbourhood is offered an alternative simultaneously.
- Target group-based approach for existing residents near station: An exploration of suburban station areas shows that the station areas often have many homes with a more bigger single-family homes. A target group-based approach may help to encourage current residents of station areas to use public transport or active modes more frequently instead of cars. Further research is needed on this group based approach.

Follow-up research on why the car is used
This research is mainly focused on analysing travel behaviour based on area types, revealing some relationships. However it does not explain why residents in suburban areas use cars more often. Therefore, it is recommended to conduct follow-up research into underlying reasons for higher car use. Possible research topics include:

- People’s attitude towards cars, bicycle and public transport.
- The origin and destination of trips. To better understand about the routes being taken.
- Other factors that may influence car use, such as comfort, safety, ownership/lease or costs.

Detailed data is needed for this follow-up research. It is recommended to obtain this data through existing traffic surveys, such as the ODiN survey, where additional work can be purchased, supplemented with extra questions regarding residents attitudes towards cars.

Working with business to promote sustainable mobility
Home-work trips make up a large part of daily mobility. This study shows that a part of the short home-work trips are made by car, while in theory these trips could be made by bicycle. However 2/3 of companies, especially small and medium-sized enterprises more present in suburban areas (VNG, 2024), do not offer bicycle or public transport allowances to their employees (CNV, 2024). Therefore, the following policies are recommended:

- Encourage bicycle and public transport allowances for employees: The province could encourage companies to introduce bicycle and public transport allowances, in combination with lobbying national and local governments.
- Explore options for mobility requirements in permit granting: In the case of new business developments, measures could be explored to enforce mobility measures. For example, by making bicycle and public transport allowances mandatory or a public transport subscription accompanying every lease car. Of course, exceptions for certain companies should also be explored.
- Introduce a mobility score for companies: Following the example of the CO2 performance ladder, possibilities could be explored to introduce a mobility score to encourage more sustainable mobility choices. Companies with a higher score could be prioritised for provincial tenders.

The recommendations for future research can help to deepen and strengthen the current research. The policy recommendations do not only provide starting points for further research, but also concrete measures to make mobility in suburban areas less car-oriented. This is essential to turn around the trend of increasing car ownership in suburban areas.

References

Abrahamse, J. E. (2019). *Opkomst en ontwikkeling van de bloemkoolwijk* (tech. rep.). Rijksdienst voor het Cultureel Erfgoed. www.cultureelerfgoed.nl

Autoriteit Consument en Markt. (2021, January). Vervoersmonitor 2019. https://www.acm.nl/nl/publicaties/vervoersmonitor-2019

Bagley, M. N., & Mokhtarian, P. L. (2002). The impact of residential neighborhood type on travel behavior: A structural equations modeling approach. *The Annals of Regional Science*, *36*, 279–297. https://doi.org/10.1007/s001680200083

Bakk, Z., Tekle, F. B., & Vermunt, J. K. (2013). Estimating the association between latent class membership and external variables using bias-adjusted three-step approaches. *Sociological Methodology*, *43*, 272–311. https://doi.org/10.1177/0081175012470644

Bakk, Z., & Vermunt, J. K. (2015). Step-3 Tutorial 1: Step-3 models with covariates, distal outcomes, and multiple latent variables. https://www.statisticalinnovations.com/wp-content/uploads/LGtutorial.Step3_.1.pdf

Banister, D., & Berechman, Y. (2001). Transport investment and the promotion of economic growth. *Journal of Transport Geography*, *9*(3), 209–218. https://doi.org/https://doi.org/10.1016/S0966-6923(01)00013-8

Batty, M. (2001). Polynucleated urban landscapes. *Urban studies*, *38*(4), 635–655. https://doi.org/10.1080/00420980120035268

Boarnet, M., & Crane, R. (2001). The influence of land use on travel behavior: Specification and estimation strategies. *Transportation Research Part A: Policy and Practice*, *35*, 823–845. https://doi.org/10.1016/S0965-8564(00)00019-7

Bolck, A., Croon, M., & Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis*, *12*, 3–27.

Boomkens, R., Van Delden, B., De Klerk, L., Musterd, S., Yap, H., & Van Der Wouden, R. (1997). Stad zonder horizon; stadspolitiek en stedelijke ontwikkeling in Nederland. -.

Boussauw, K., Allaert, G., & Witlox, F. (2012). Stedelijke subcentra en korte verplaatsingen: Is er een verband? het geval van de lagere scholen in vlaanderen. https://www.cvs-congres.nl/cvspdfdocs/cvs12_6.pdf

Bramer, W., de Jonge, G., Rethlefsen, M., Mast, F., & Kleijnen, J. (2018). A systematic approach to searching: An efficient and complete method to develop literature searches. *Journal of the Medical Library Association*, *106*, 531–541. https://doi.org/10.5195/jmla.2018.283

Brownson, R. C., Hoehner, C. M., Day, K., Forsyth, A., & Sallis, J. F. (2009). Measuring the built environment for physical activity: State of the science. *American Journal of Preventive Medicine*, *36*, S99–S123.e12. https://doi.org/https://doi.org/10.1016/j.amepre.2009.01.005

CBS. (2019). *Onderzoek Onderweg in Nederland - ODiN 2018*. https://doi.org/10.17026/dans-xpv-mwpg

CBS. (2020a). Kerncijfers per postcode. https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/gegevens-per-postcode

CBS. (2020b). *Onderzoek Onderweg in Nederland - ODiN 2019*. https://doi.org/10.17026/dans-xpv-mwpg

CBS. (2022a). Kerncijfers per postcode. https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/gegevens-per-postcode

CBS. (2022b, July). Literatuur - PBL/CBS Regionale bevolkings- en huishoudensprognose 2022–2050. https://longreads.cbs.nl/regionale-prognose-2022/literatuur/

CBS. (2023a). Regionale kerncijfers Nederland. https://opendata.cbs.nl/statline/?dl=16D63#/CBS/nl/dataset/70072ned/table

CBS. (2023b, May). Nieuwe auto's en jaarkilometrage per PC4, 2019 en 2021. https://www.cbs.nl/nl-nl/maatwerk/2023/18/nieuwe-auto-s-en-jaarkilometrage-per-pc4-2019-en-2021

CBS. (2024a). *Onderzoek Onderweg in Nederland - ODiN 2023*. https://doi.org/10.17026/dans-xpv-mwpg

CBS. (2024b). Stedelijkheid (van een gebied). https://www.cbs.nl/nl-nl/onze-diensten/methoden/begrippen/stedelijkheid%E2%80%93van-een-gebied%E2%80%93#:~:text=%2D%20sterk%20stedelijk%3A%20gemiddelde%20oad%20van,dan%20500%20adressen%20per%20km2.

Cervero, R., & Duncan, M. (2006). 'which reduces vehicle travel more: Jobs-housing balance or retail-housing mixing? *Journal of the American Planning Association © American Planning Association*, *72*. https://doi.org/10.1080/01944360608976767

Cervero, R., & Radisch, C. (1996). Travel choices in pedestrian versus automobile oriented neighborhoods. *Transport Policy*, *3*(3), 127–141. https://doi.org/https://doi.org/10.1016/0967-070X(96)00016-9

Chen, C., Gong, H., & Paaswell, R. E. (2007). Role of the built environment on mode choice decisions: additional evidence on the impact of density. *Transportation*, *35*(3), 285–299. https://doi.org/10.1007/s11116-007-9153-5

Chen, C., & McKnight, C. E. (2007). Does the built environment make a difference? additional evidence from the daily activity and travel behavior of homemakers living in new york city and suburbs. *Journal of Transport Geography*, *15*(5), 380–395. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2006.11.010

Christiansen, H., & Baescu, O. (2022). *The danish national travel survey: Annual statistical report for copenhagen area for 2021*.

Christiansen, P., Engebretsen, Ø., Fearnley, N., & Hanssen, J. U. (2017). Parking facilities and the built environment: Impacts on travel behaviour. *Transportation Research Part A: Policy and Practice*, *95*, 198–206. https://doi.org/https://doi.org/10.1016/j.tra.2016.10.025

CNV. (2024, July). CNV-onderzoek: meerderheid werknemers krijgt geen fietsvergoeding | CNV. https://www.cnv.nl/nieuws/cnv-onderzoek-meerderheid-werknemers-krijgt-geen-fietsvergoeding/

Crane, R. (1995). On form versus function: Will the "new urbanism" reduce traffic or increase it?

Cropper, M., Vinha, K., Bento, A., & Mobarak, A. (2005). The effects of urban spatial structure on travel demand in the united states. *The Review of Economics and Statistics*, *87*, 466–478. https://doi.org/10.1162/0034653054638292

Dandy, P. (1997). Urban density and car and bus use in edinburgh. www.witpress.com

Dargay, J., & Hanly, M. (2004). Land use and mobility.

Dembski, S. (2015). Structure and imagination of changing cities: Manchester, liverpool and the spatial in-between. *Urban Studies*, *52*. https://doi.org/10.1177/0042098014539021

Department of transport. (2024, August). Region and Rural-Urban Classification. https://www.gov.uk/government/statistical-data-sets/nts99-travel-by-region-and-area-type-of-residence#mode-of-transport

Dieleman, F. M., Dijst, M., & Burghouwt, G. (2002). Urban form and travel behaviour: Micro-level household attributes and residential context. *Urban Studies*, *39*(3), 507–527. https://doi.org/10.1080/00420980220112801

Ding, C., Wang, D., Liu, C., Zhang, Y., & Yang, J. (2017). Exploring the influence of built environment on travel mode choice considering the mediating effects of car ownership and travel distance. *Transportation Research Part A: Policy and Practice*, *100*, 65–80. https://doi.org/https://doi.org/10.1016/j.tra.2017.04.008

Ewing, R., & Cervero, R. (2001). Travel and the built environment: A synthesis. *Transportation Research Record*, *1780*. https://doi.org/10.3141/1780-10

Ewing, R., & Cervero, R. (2010). Travel and the built environment. *Journal of the American Planning Association*, *76*, 265–294. https://doi.org/10.1080/01944361003766766

Flynt, A., & Dean, N. (2016). A survey of popular r packages for cluster analysis. *Journal of Educational and Behavioral Statistics*, *41*(2), 205–225. https://doi.org/10.3102/1076998616631743

Forsyth, A. (2012). Defining suburbs. *Journal of Planning Literature*, *27*, 270–281. https://doi.org/10.1177/0885412212448101

Frank, L. D., & Engelke, P. O. (2001). The built environment and human activity patterns: Exploring the impacts of urban form on public health. *Journal of Planning Literature*, *16*, 202–218. https://doi.org/10.1177/08854120122093339

Friedman, B., Gordon, S. R., & Peers, J. B. (1994). Effect of neotraditional neighborhood design on travel characteristics. *Transportation Research Record*. https://api.semanticscholar.org/CorpusID:107060651

GHSL - Global Human Settlement Layer. (2016, July). Global Human Settlement - Urban centres database 2018 visualisation - European Commission. https://human-settlement.emergency.copernicus.eu/ucdb2018visual.php

Gruyter, C. D., Truong, L. T., de Jong, G., & Foster, S. (2024). Determinants of zero-car and car-owning apartment households. *Transportation*. https://doi.org/10.1007/s11116-024-10467-8

Hamers, D. (2003, November). *Tijd voor suburbia: De amerikaanse buitenwijk in wetenschap en literatuur*.

Handy, S. L., Boarnet, M. G., Ewing, R., & Killingsworth, R. E. (2002). How the built environment affects physical activity: Views from urban planning. *American Journal of Preventive Medicine*, *23*, 64–73. https://doi.org/https://doi.org/10.1016/S0749-3797(02)00475-0

Harris, R., & Larkham, P. (2004). *Changing suburbs: Foundation, form and function*. Taylor & Francis. https://books.google.nl/books?id=ExYEE-XQvcUC

Hassler, U., & Kohler, N. (2014). Resilience in the built environment. *Building Research Information*, *42*, 119–129. https://doi.org/10.1080/09613218.2014.873593

Izenman, A. J. (2008, August). *Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning*. https://ci.nii.ac.jp/ncid/BB19457782

Jalali, S., & Wohlin, C. (2012). Systematic literature studies: Database searches vs. backward snowballing. *Proceedings of the 2012 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement*, 29–38. https://doi.org/10.1145/2372251.2372257

James Cook University. (2023, February). Library guides: Literature reviews: Identify keywords. https://libguides.jcu.edu.au/litreview/keywords

Kaplan, D. E. (2004, January). *The SAGE Handbook of Quantitative Methodology for the Social Sciences*. https://doi.org/10.4135/9781412986311

Khattak, A., & Rodriguez, D. (2004). The impact of neo-traditional neighborhood developments on travel behavior.

KIM. (2021). MPN 2021. https://www.mpndata.nl/study-units/view/13

KiM. (2022). Het wijdverbreide autobezit in nederland. https://www.kimnet.nl/mobiliteitspanel-nederland/publicaties/publicaties/2022/02/22/het-wijdverbreide-autobezit-in-nederland

Knudsen, M. A. (2014, May). *Danish long distance travel A study of Danish travel behaviour and the role of infrequent travel activities*. https://orbit.dtu.dk/en/publications/danish-long-distance-travela-study-of-danish-travel-behaviour-and

Kockelman, K. (1997). Travel behavior as function of accessibility, land use mixing, and land use balance: Evidence from san francisco bay area. *Transportation Research Record*, *1607*, 116–125. https://doi.org/10.3141/1607-16

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/https://doi.org/10.1016/j.jth.2018.11.006

Krygsman, S., Dijst, M., & Arentze, T. (2004). Multimodal public transport: An analysis of travel time elements and the interconnectivity ratio. *Transport Policy*, *11*(3), 265–275. https://doi.org/https://doi.org/10.1016/j.tranpol.2003.12.001

Kubeš, J., & Ouředníček, M. (2022). Functional types of suburban settlements around two differently sized czech cities. *Cities*, *127*, 103742. https://doi.org/https://doi.org/10.1016/j.cities.2022.103742

Lee, Y., Circella, G., Mokhtarian, P., & Guhathakurta, S. (2019, June). *Are millennials more multimodal? a latent-class cluster analysis with attitudes and preferences among millennial and generation x commuters in california*.

Lenormand, M., Louail, T., Ros, O. G. C., Picornell, M., Herranz, R., Arias, J., Barthelemy, M., Miguel, M., & Ramasco, J. J. (2014). Influence of sociodemographics on human mobility. *Scientific reports*, *5*. https://doi.org/10.1038/srep10075

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

LISA. (2022). Landelijk Informatiesysteem van Arbeidsplaatsen.

Liu, L., & Alain, L. (2014). Transport and land use interaction: A french case of suburban development in the lille metropolitan area (lma). *Transportation Research Procedia*, *4*, 120–139. https://doi.org/10.1016/j.trpro.2014.11.011

Lunke, E., & Engebretsen, Ø. (2023). Public transport use on trip chains: Exploring various mode choice determinants. *Findings*. https://doi.org/10.32866/001c.74112

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

Martensen, H., & Arendsen, K. (2024a). *Auto atlas, achtergrondrapport*. Kennisinstituut voor Mobiliteitsbeleid (KiM).

Martensen, H., & Arendsen, K. (2024b, August). *Atlas van de auto* (tech. rep.). https://www.kimnet.nl/publicaties/publicaties/2024/06/03/atlas-van-de-auto

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. (2022, October). Provinciale woningbouwafspraken 2022-2030 aantallen per provincie. https://www.rijksoverheid.nl/documenten/rapporten/2022/10/13/provinciale-woningbouwafspraken-2022-2030-aantallen-per-provincie

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/https://doi.org/10.1016/j.tra.2015.11.001

Morgan, G. B. (2015). Mixed mode latent class analysis: An examination of fit index performance for classification. *Structural Equation Modeling: A Multidisciplinary Journal*, *22*(1), 76–86. https://doi.org/10.1080/10705511.2014.935751

Nylund-Gibson, K., & Choi, A. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science*, *4*, 440–461. https://doi.org/10.1037/tps0000176

Peng, Z.-R. (1997). The jobs-housing balance and urban commuting. *Urban Studies*, *34*(8), 1215–1235. https://doi.org/10.1080/0042098975600

Planbureau voor de Leefomgeving. (2006). Vele steden maken nog geen randstad. https://www.pbl.nl/sites/default/files/downloads/Vele_steden_maken_nog_geen_randstad.pdf

Provincie Zuid-Holland. (2011, April). *Hoofdlijnenakkoord 2011 - 2015*.

Provincie Zuid-Holland. (2015, May). *Hoofdlijnenakkoord 2015-2019*. https://www.zuid-holland.nl/publish/pages/10694/zuid-hollandslimmerschonerensterkerhoofdlijnenakkoord2015-2019.pdf

Provincie Zuid-Holland. (2023a). Databank Zuid-Holland. https://www.databankzh.nl/database/data/data.php?nr=558&visual=table

Provincie Zuid-Holland. (2023b). Motorvoertuigenbezit Provincie Zuid-Holland. https://www.zuid-holland.nl/portfolio_page/motorvoertuigenbezit/

Provincie Zuid-Holland. (2023c). Regionale realisatieagenda's wonen. https://www.zuid-holland.nl/onderwerpen/ruimte/versnellen-woningbouw/regionale-realisatieagenda-wonen/

Provincie Zuid-Holland. (2023d, November). Motorvoertuigenbezit - staat van Zuid-Holland. https://staatvan.zuid-holland.nl/portfolio_page/motorvoertuigenbezit/

Pucher, J., & Buehler, R. (2017). Transport reviews cycling towards a more sustainable transport future. *Transport Reviews*, *37*, 384–389.

Reboussin, B. A., Ip, E. H., & Wolfson, M. (2008). Locally dependent latent class models with covariates: an application to under□age drinking in the USA. *Journal of the Royal Statistical Society Series A*, *171*(4), 877–897. https://doi.org/10.1111/j.1467-985X.2008.

ruimtemettoekomst.nl. (2013, May). Ruimte met toekomst - suburbaan gebied. http://www.ruimtexmilieu.nl/wiki/gebiedstypering/suburbaan-gebied

Schwanen, T., Dieleman, F., & Dijst, M. (2004). The impact of metropolitan structure on commute behavior in the netherlands: A multilevel approach. *Growth and Change*, *35*, 304–333. https://doi.org/10.1111/j.1468-2257.2004.00251.x

Son, S., Khattak, A., Wang, X., Agnello, P., & Chen, J.-Y. (2013). Quantifying key errors in household travel surveys. *Transportation Research Record: Journal of the Transportation Research Board*, *2354*, 9–18. https://doi.org/10.3141/2354-02

Spears, S., Boarnet, M. G., Handy, S., & Rodier, C. (2014). Impacts of Land-Use Mix on Passenger Vehicle Use and Greenhouse Gas Emissions. https://ww2.arb.ca.gov/sites/default/files/2020-06/Impacts_of_Land-Use_Mix_on_Passenger_Vehicle_Use_and_Greenhouse_Gas_Emissions_Technical_Background_Document_0.pdf

Stead, D., & Marshall, S. (2001). The relationships between urban form and travel patterns. an international review and evaluation. *Stead, D. and Marshall, S. (2001) The relationships between urban form and travel patterns: an international review and evaluation. European Journal of Transport and Infrastructure Research, 1 (2). pp. 113-141. ISSN 15677141*, *1*. https://doi.org/10.18757/ejtir.2001.1.2.3497

Stoker, P., & Ewing, R. (2014). Job–worker balance and income match in the united states. *Housing Policy Debate*, *24*, 485–497. https://doi.org/10.1080/10511482.2013.852604

Ton, D., Zomer, L.-B., Schneider, F., Hoogendoorn-Lanser, S., Duives, D., Cats, O., & Hoogendoorn, S. (2020). Latent classes of daily mobility patterns: The relationship with attitudes towards modes. *Transportation*, *47*, 1843–1866. https://doi.org/10.1007/s11116-019-09975-9

Van Acker, V., & Witlox, F. (2009). Car ownership as a mediating variable in car travel behaviour research using a structural equation modelling approach to identify its dual relationship. *Journal of Transport Geography*, *18*(1), 65–74. https://doi.org/10.1016/j.jtrangeo.2009.05.006

Van Acker, V., & Witlox, F. (2010). Car ownership as a mediating variable in car travel behaviour research using a structural equation modelling approach to identify its dual relationship. *Journal of Transport Geography*, *18*(1), 65–74. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2009.05.006

van Ham, M. (2012). Economics of housing choice. In S. Smith, M. Elsinga, L. Fox-O'Mahony, S. Ong, & S. Wachter (Eds.), *The international encyclopedia of housing and home* (pp. 42–46). Elsevier.

van Wee, B., De Vos, J., & Maat, K. (2019). Impacts of the built environment and travel behaviour on attitudes: Theories underpinning the reverse causality hypothesis. *Journal of Transport Geography*, *80*, 102540. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2019.102540

Vandenbergh, J. (2024). *Buurtvoorzieningen. Onderzoek naar de economische dynamiek van het voorzieningenaanbod* (tech. rep.).

van Wee, B. (2009). Self-selection: A key to a better understanding of location choices, travel behaviour and transport externalities? *Transport Reviews*, *29*(3), 279–292. https://doi.org/10.1080/01441640902752961

Vega, A., & Reynolds-Feighan, A. (2009a). A methodological framework for the study of residential location and travel-to-work mode choice under central and suburban employment destination patterns. *Transportation Research Part A: Policy and Practice*, *43*, 401–419. https://doi.org/https://doi.org/10.1016/j.tra.2008.11.011

Vega, A., & Reynolds-Feighan, A. (2009b). A methodological framework for the study of residential location and travel-to-work mode choice under central and suburban employment destination patterns. *Transportation Research Part A: Policy and Practice*, *43*(4), 401–419. https://doi.org/https://doi.org/10.1016/j.tra.2008.11.011

Vermunt, J. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Psycho-oncology - PSYCHO-ONCOL*, *18*. https://doi.org/10.2307/25792024

Vermunt, J. K., & Magidson, J. (2016). Technical guide for latent gold 5.1: Basic, advanced, and syntax 1. https://api.semanticscholar.org/CorpusID:64107953

Vermunt, J. K., & Magidson, J. (2020). How to Perform Three-Step Latent Class Analysis in the Presence of Measurement Non-Invariance or Differential Item Functioning. *Structural Equation Modeling A Multidisciplinary Journal*, *28*(3), 356–364. https://doi.org/10.1080/10705511.2020.1818084

VNG. (2024). Vereniging Nederlandse Gemeenten, Dashboard - Werk en inkomen. https://www.waarstaatjegemeente.nl/mosaic/dashboard/werk-en-inkomen

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the Association for Information Systems*, *37*, 205–224. https://doi.org/10.17705/1cais.03709

Wang, X., Yan, X., Zhao, X., & Cao, Z. (2022). Identifying latent shared mobility preference segments in low-income communities: Ride-hailing, fixed-route bus, and mobility-on-demand transit. *Travel Behaviour and Society*, *26*, 134–142. https://doi.org/https://doi.org/10.1016/j.tbs.2021.09.011

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

Wiersma, J. (2020). Commuting patterns and car dependency in urban regions. *Journal of Transport Geography*, *84*. https://doi.org/10.1016/j.jtrangeo.2020.102700

www.prettigparkeren.nl. (2024). Prettig Parkeren. https://www.prettigparkeren.nl/kaart/

A

Car ownership

A.1. Car ownership table

This appendix shows the high increase in car ownership in the province of Zuid-Holland from 2001 to 2021. The increases are mainly concentrated in the more rural and suburban area's of Zuid-Holland.

GEMEENTE	2001	2006	2011	2016	2021	increase 5y	increase 10y	increase 20y
Teylingen	404,8	424,7	444,9	484,6	534,5	10,3	20,1	32,0
Katwijk	339,2	359,5	379,8	383,6	422,5	10,1	11,2	24,6
Sliedrecht	380,4	407,0	439,3	435,8	474,5	8,9	8,0	24,7
Dordrecht	363,7	399,9	437,7	459,9	500,2	8,8	14,3	37,5
Gorinchem	378,7	399,9	424,9	431,2	464,0	7,6	9,2	22,5
Rotterdam	318,1	331,1	346,2	343,2	368,9	7,5	6,6	16,0
Westvoorne	471,5	510,9	556,9	573,3	615,6	7,4	10,5	30,6
Molenlanden	404,2	432,9	475,0	497,6	532,9	7,1	12,2	31,8
Krimpenerwaard	385,3	414,5	449,9	471,6	503,8	6,8	12,0	30,8
Alblasserdam	361,2	397,5	431,7	440,9	469,4	6,5	8,7	30,0
Goeree-Overflakkee	406,3	445,7	495,1	515,9	547,1	6,0	10,5	34,7
Noordwijk	402,5	433,3	467,0	482,9	511,0	5,8	9,4	26,9
Vlaardingen	359,7	381,1	402,8	409,4	432,8	5,7	7,4	20,3
Brielle	446,4	474,0	514,1	535,0	565,5	5,7	10,0	26,7
Lisse	435,1	451,3	466,1	494,2	521,0	5,4	11,8	19,7
Nieuwkoop	415,9	446,2	486,0	518,8	546,7	5,4	12,5	31,5
Den Haag	328,2	350,0	356,8	345,4	363,9	5,3	2,0	10,9
Nissewaard	385,7	419,5	458,6	464,1	488,8	5,3	6,6	26,7
Kaag en Braassem	414,8	448,7	487,8	507,4	534,1	5,3	9,5	28,7
Bodegraven-Reeuwijk	394,3	424,7	453,4	473,9	498,5	5,2	10,0	26,4
Capelle aan den IJssel	392,3	425,0	441,2	442,1	464,7	5,1	5,3	18,5
Krimpen aan den IJssel	388,3	416,5	440,1	446,5	468,1	4,8	6,4	20,5
Zwijndrecht	386,9	417,3	448,6	467,1	488,3	4,5	8,8	26,2
Barendrecht	420,9	441,7	462,6	466,2	487,2	4,5	5,3	15,8
Pijnacker-Nootdorp	391,9	410,3	427,3	436,5	455,9	4,4	6,7	16,3
Westland	414,3	453,1	487,9	491,0	512,7	4,4	5,1	23,8
Zoetermeer	377,5	412,1	439,0	444,8	464,1	4,4	5,7	23,0
Hoeksche Waard	425,3	457,5	496,4	521,0	543,7	4,4	9,5	27,8
Hellevoetsluis	391,3	435,3	478,7	503,4	524,4	4,2	9,5	34,0
Maassluis	367,7	401,8	429,8	441,7	460,0	4,2	7,0	25,1
Zuidplas	400,2	434,8	477,7	490,0	510,1	4,1	6,8	27,5
Papendrecht	410,3	424,7	461,5	473,2	491,3	3,8	6,4	19,7
Hillegom	428,8	460,5	487,3	507,8	526,8	3,7	8,1	22,9
Waddinxveen	411,2	435,7	470,8	489,4	506,9	3,6	7,7	23,3
Ridderkerk	412,9	448,7	478,0	487,6	504,4	3,4	5,5	22,1
Leiden	375,3	363,6	376,7	313,3	323,9	3,4	-14,0	-13,7
Rijswijk	663,9	528,8	590,9	536,6	554,2	3,3	-6,2	-16,5
Delft	322,8	320,9	328,5	316,1	326,2	3,2	-0,7	1,1
Hardinxveld-Giessendam	384,0	409,3	440,6	464,4	479,0	3,1	8,7	24,7
Hendrik-Ido-Ambacht	398,4	425,0	446,6	456,6	470,3	3,0	5,3	18,1
Albrandswaard	413,4	440,6	468,4	484,3	498,5	2,9	6,4	20,6
Gouda	863,6	344,1	363,9	382,7	393,1	2,7	8,0	-54,5
Lansingerland	431,0	458,7	496,7	487,3	500,4	2,7	0,8	16,1
Midden-Delfland	414,6	425,0	454,7	471,7	481,8	2,1	6,0	16,2
Voorschoten	417,0	434,4	455,4	450,1	457,1	1,6	0,4	9,6
Zoeterwoude	401,4	427,6	460,6	476,3	477,0	0,1	3,5	18,8
Schiedam	339,0	358,7	381,1	374,5	374,5	0,0	-1,7	10,5
Oegstgeest	390,2	407,9	419,5	435,5	432,7	-0,6	3,2	10,9
Leiderdorp	455,0	430,4	448,3	480,9	476,9	-0,8	6,4	4,8
Leidschendam-Voorburg	414,6	435,7	462,3	470,8	456,4	-3,1	-1,3	10,1
Wassenaar	500,5	512,7	518,8	545,8	509,2	-6,7	-1,9	1,7
Alphen aan den Rijn	394,1	480,3	506,2	523,3	485,2	-7,3	-4,1	23,1

Figure A.1: Table of car ownership per municipality per 1000 inhabitants in 2001,2006,2011,2016 and 2021. (Note: Schiedam ownership is not reported correctly in 2021 and Gouda data also appears to be not correct in 2001) (Provincie Zuid-Holland, 2023c)

B

Expert panel questionnaire

B.1. Results from survey

This survey is conducted on 15 mobility experts working at the Province of Zuid-Holland. It should be noted that this are personal opinions from the mobility experts, based on their residential location.

	Mean	Median	SD	Min	Max	Used in formula
Supermarket	3,9	3	2,3	2	10	3
Daily groceries	14,1	11	10,0	2	35	12
Primairy School	2,1	2	1,1	0	5	2
High School	2,2	2	0,6	1	3	3
Café	1,8	2	1,5	0	5	2
Cafetaria	2,1	2	1,4	0	4	1
Hotel	2,3	2	1,9	0	5	2
restaurant	9,8	9	8,0	2	30	4
Attraction	2,8	3	1,9	0	6	2
Cinema	1,6	2	0,9	0	3	2
Musea	2,7	2,5	2,3	0	8	2
Hospital (with beds)	1,5	1	0,7	1	3	1

Table B.1: Values from the experts review, right column is showing the value used imn the formula.

Figure B.1: Question answers: Are there certain amenities that you feel are very important to your sense of proximity to amenities?

Figure B.2: Question answers: Are there certain amenities that you feel are less important to your sense of proximity to amenities?

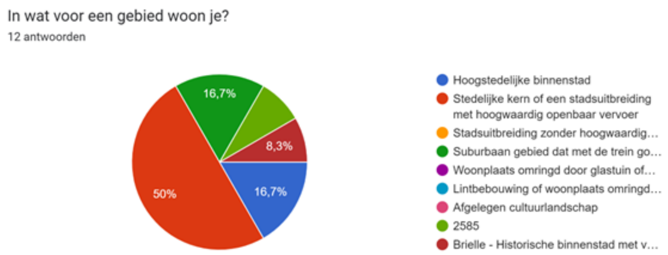


Figure B.3: Question answers: What sort of area do you live?

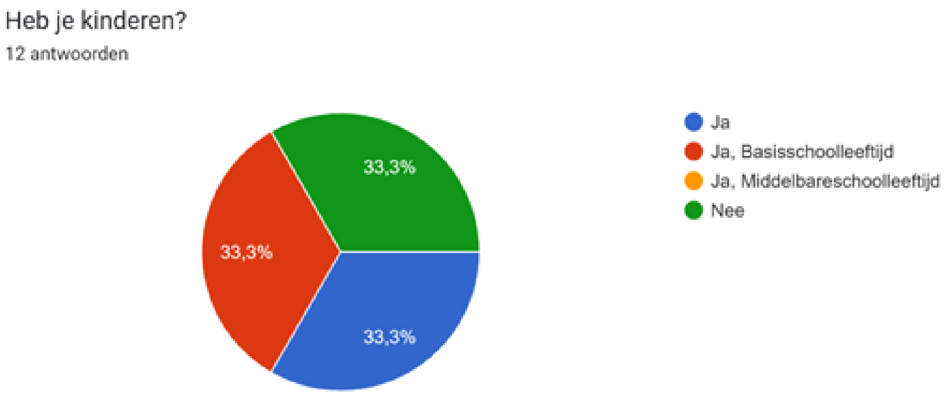


Figure B.4: Question answers: Do you've Childeren

B.2. Survey questions and design

Nabijheidscore voorzieningen.

Uit CBS data kunnen we voor elke wijk achterhalen hoeveel voorzieningen er aanwezig zijn. Voor Zuid-Holland heb ik dit in de onderstaande tabel gezet. Je ziet dat er doorgaan 7 supermarkten zijn binnen straal van 3 kilometer.

Deze aantallen zijn wel anders dan wat de "basis nabijheid" is, dat is namelijk het aantal voorzieningen dat aanwezig moet zijn zodat je voldoende keuze mogelijkheden hebt. Deze vragenlijst gaat mij daar meer inzicht in geven.

De vragenlijst is opgedeeld in 6 secties, bij de eerste 4 zal je gevraagd worden even naar je eigen wijk te kijken op een speciale link. Lukt dit toch niet, laat het even weten, dan help ik je erbij.
* 1 kilometer
* 3 kilometer
* 5 kilometer
* 10 kilometer
* waardering
* afronding

Alvast bedankt voor het invullen!

m.r.klootwijk@gmail.com Ander account

Niet gedeeld

	Aantal (Mediaan)	Binnen een afstand van:
Basisschool	1.8	1 km
Café	0.8	1km
Cafetaria	2.4	1km
Supermarkt	7.1	3 km
Dagelijkse levensmiddelen	25.9	3 km
Middelbare school	2.8	3 km
Restaurant	21.3	3 km
Hotel	3.8	5 km
Bioscoop	0.9	5 km
Musea	1.1	5 km
Ziekenhuis (met opname langer dan 24 uur)	0.9	5 km
Attractie	3	10 km

Volgende

Formulier wissen

Voorzieningen binnen een afstand van 1 kilometer

Ga naar deze link en vul je eigen adres in en een straal van **1 kilometer**. Kijk goed naar de cirkel en maak een inschatting hoeveel voorzieningen jij binnen dit gebied zou willen hebben voor "net genoeg keuze". Dit kan dus meer of minder zijn dan er in jouw buurt aanwezig is.

	Aantal (Mediaan)	**Binnen een afstand van:**
Basisschool	1.8	1 km
Café	0.8	1km
Cafetaria	2.4	1km

Hoeveel **basisscholen** moeten er binnen 1 kilometer van je woning zijn zodat jij keuzevrijheid ervaart? *

1

Hoeveel **cafe's** en dergelijke moeten er binnen 1 kilometer van je woning zijn zodat jij keuzevrijheid ervaart? *

denk aan: Café, koffiehuis, coffeeshop, discotheek, seks/nachtclub en partycentrum.

1

Hoeveel **cafetaria's** en dergelijke moeten er binnen 1 kilometer van je woning zijn zodat jij keuzevrijheid ervaart? *

denk aan: Snackbar, Fastfoodrestaurant, grillroom/shoarmazaak, lunchroom, pannenkoekenhuis en ijssalon.

1

Voorzieningen binnen een afstand van 3 kilometer

Ga naar deze link en vul je eigen adres in en een straal van **3 kilometer**. Kijk goed naar de cirkel en maak een inschatting hoeveel voorzieningen jij binnen dit gebied zou willen hebben voor "net genoeg keuze". Dit kan dus meer of minder zijn dan er in jouw buurt aanwezig is.

	Aantal (Mediaan)	Binnen een afstand van:
Supermarkt	7.1	3 km
Overige dagelijkse levensmiddelen	25.9	3 km
Middelbare school	2.8	3 km
Restaurant	21.3	3 km

Hoeveel **supermarkten** moeten er binnen 3 kilometer van je woning zijn zodat **jij** keuzevrijheid ervaart? *

1

Hoeveel winkels voor **dagelijkse levensmiddelen** moeten er binnen 3 kilometer van je woning zijn zodat **jij** keuzevrijheid ervaart? *

Denk aan: Groenteboer, bakker, vlaaienwinkel, toko, chocoladewinkel, koffie/theewinkel, delicatessenwinkel, kaaswinkel, mini supermarkt, notenwinkel, poelier, reformwinkel, slagerij, slijterij, tabakswinkel, visboer, zoetwarenwinkel, nachtwinkel, diepvriesartikelenwinkel, wijnwinkel en ziekenhuiswinkel.

1

Hoeveel **middelbare scholen** moeten er binnen 3 kilometer van je woning zijn zodat jij keuzevrijheid ervaart? *

1

Hoeveel **restaurants** moeten er binnen 3 kilometer van je woning zijn zodat jij keuzevrijheid ervaart? *

1

Voorzieningen binnen een afstand van 5 kilometer

Ga naar deze link en vul je eigen adres in en een straal van **5 kilometer**. Kijk goed naar de cirkel en maak een inschatting hoeveel voorzieningen jij binnen dit gebied zou willen hebben voor "net genoeg keuze". Dit kan dus meer of minder zijn dan er in jouw buurt aanwezig is.

	Aantal (Mediaan)	**Binnen een afstand van:**
Hotel	3.8	5 km
Bioscoop	0.9	5 km
Musea	1.1	5 km
Ziekenhuis (met opname langer dan 24 uur)	0.9	5 km

Hoeveel **hotels** moeten er binnen 5 kilometer van je woning zijn zodat jij (of jouw bezoek) keuzevrijheid ervaart? *

1

Hoeveel **bioscopen** moeten er binnen 5 kilometer van je woning zijn zodat jij keuzevrijheid ervaart? *

1

Hoeveel **musea** moeten er binnen 5 kilometer van je woning zijn zodat jij keuzevrijheid ervaart? *

1

Hoeveel **ziekenhuizen (met opname)** moeten er binnen 5 kilometer van je woning zijn zodat jij keuzevrijheid ervaart? *

1

Voorzieningen binnen een afstand van 10 kilometer

Ga naar deze link en vul je eigen adres in en een straal van **10 kilometer**. Kijk goed naar de cirkel en maak een inschatting hoeveel voorzieningen jij binnen dit gebied zou willen hebben voor "net genoeg keuze". Dit kan dus meer of minder zijn dan er in jouw buurt aanwezig is.

	Aantal (Mediaan)	**Binnen een afstand van:**
Attractie	3	10 km

Hoeveel **attracties** moeten er binnen 10 kilometer van je woning zijn zodat jij keuzevrijheid ervaart? *

denk aan: Pretpark, dierentuin en binnenspeeltuin.

1

Vorige Volgende Formulier wissen

Wegingen nabijheid

Voor het bepalen of sommige voorzieningen belangrijker zijn voor de algehele nabijheid zou ik graag willen weten welke voorzieningen jij erg belangrijk en juist niet belangrijk vindt.

	Aantal (Mediaan)	Binnen een afstand van:
Basisschool	1.8	1 km
Café	0.8	1km
Cafetaria	2.4	1km
Supermarkt	7.1	3 km
Dagelijkse levensmiddelen	25.9	3 km
Middelbare school	2.8	3 km
Restaurant	21.3	3 km
Hotel	3.8	5 km
Bioscoop	0.9	5 km
Musea	1.1	5 km
Ziekenhuis (met opname langer dan 24 uur)	0.9	5 km
Attractie	3	10 km

Zijn er bepaalde voorzieningen die jij **erg belangrijk** vindt voor jouw gevoel van nabijheid tot voorzieningen? *

meerdere opties zijn mogelijk

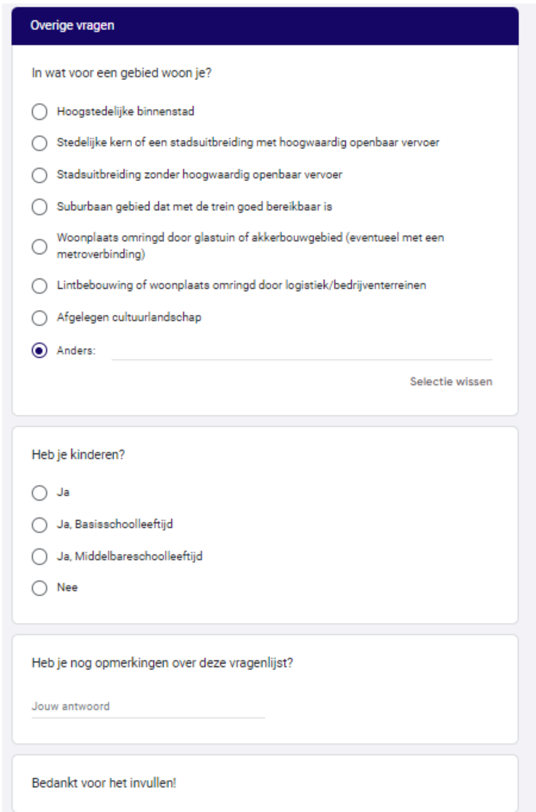
- [x] Supermarkt
- [x] Dagelijkse boodschappen
- [x] Basisschool
- [ ] Middelbare school
- [ ] Cafe's
- [ ] Cafetaria
- [ ] Hotel
- [ ] Restaurant
- [ ] Attractie
- [ ] Bioscoop
- [ ] Museum
- [ ] Ziekenhuis binnen 5 km (met mogelijkheid tot opname langer dan 24 uur)

Zijn er bepaalde voorzieningen die jij **veel minder belangrijk** vindt voor jouw gevoel van nabijheid tot voorzieningen? *

meerdere opties zijn mogelijk

- [x] Supermarkt
- [ ] Dagelijkse boodschappen
- [ ] Basisschool
- [ ] Middelbare school
- [ ] Cafe's
- [ ] Cafetaria
- [ ] Hotel
- [ ] Restaurant
- [ ] Attractie
- [ ] Bioscoop
- [ ] Museum
- [ ] Ziekenhuis binnen 5 km (met mogelijkheid tot opname langer dan 24 uur)

Vorige | Volgende | Formulier wissen



Overige vragen

In wat voor een gebied woon je?

- ◯ Hoogstedelijke binnenstad
- ◯ Stedelijke kern of een stadsuitbreiding met hoogwaardig openbaar vervoer
- ◯ Stadsuitbreiding zonder hoogwaardig openbaar vervoer
- ◯ Suburbaan gebied dat met de trein goed bereikbaar is
- ◯ Woonplaats omringd door glastuin of akkerbouwgebied (eventueel met een metroverbinding)
- ◯ Lintbebouwing of woonplaats omringd door logistiek/bedrijventerreinen
- ◯ Afgelegen cultuurlandschap
- ◉ Anders:

Selectie wissen

Heb je kinderen?

- ◯ Ja
- ◯ Ja, Basisschoolleeftijd
- ◯ Ja, Middelbareschoolleeftijd
- ◯ Nee

Heb je nog opmerkingen over deze vragenlijst?

Jouw antwoord

Bedankt voor het invullen!

C
Indicators BE on a map

The maps below display indicators of the built environment across four different areas.

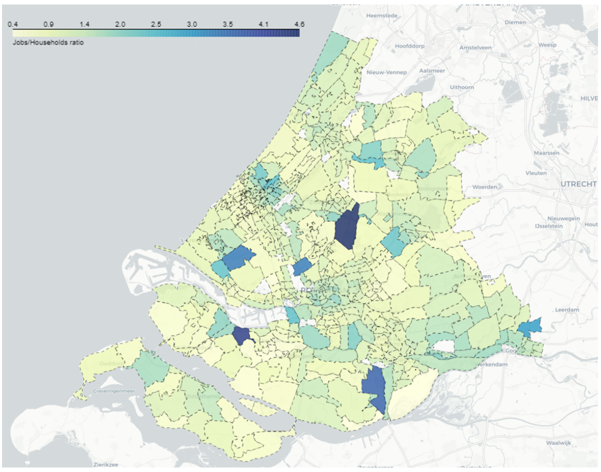


Figure C.1: Map: Job household balance

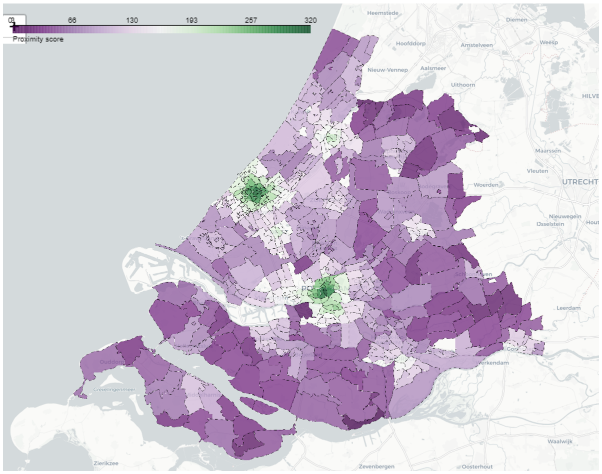


Figure C.2: Map: Proximity to facilities

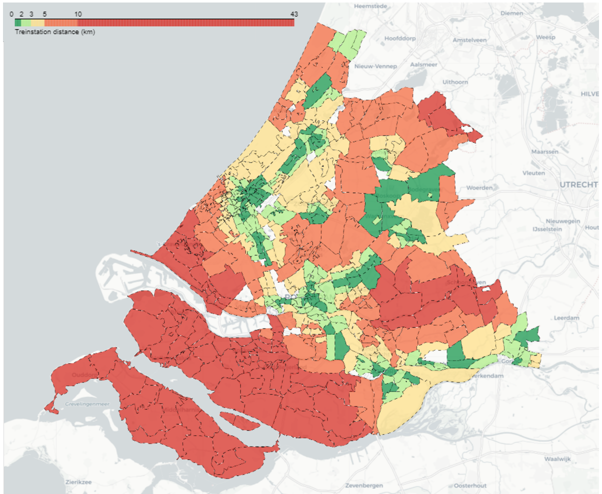


Figure C.3: Map: Distances to train stations in km

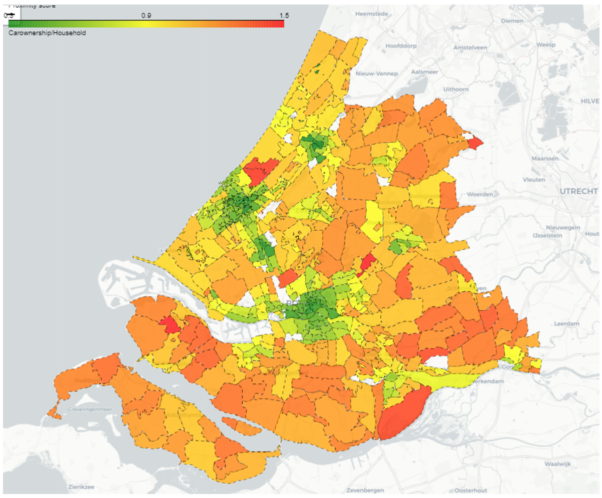


Figure C.4: Map: Car ownership

D Data sociodemograpics visualized

This appendix presents all sociodemographic data, As mentioned in the main part of the thesis all data is from ODiN 2018, 2019 and 2023.

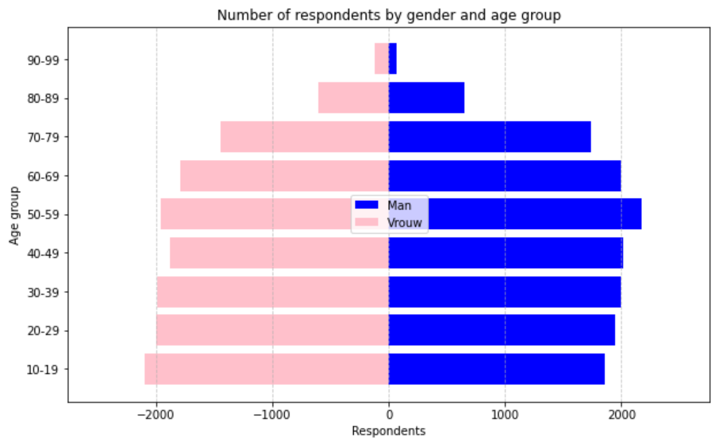


Figure D.1: Age and gender of all respondents.

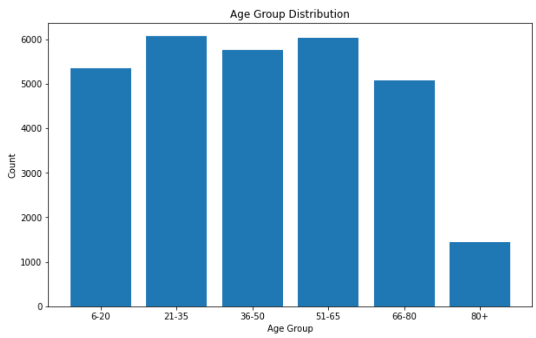


Figure D.2: Age distribution of respondents

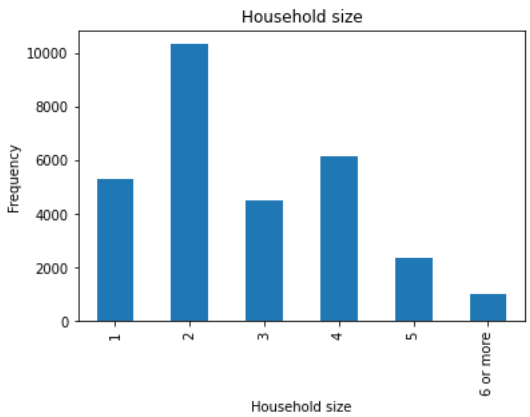


Figure D.3: Household size of respondents.

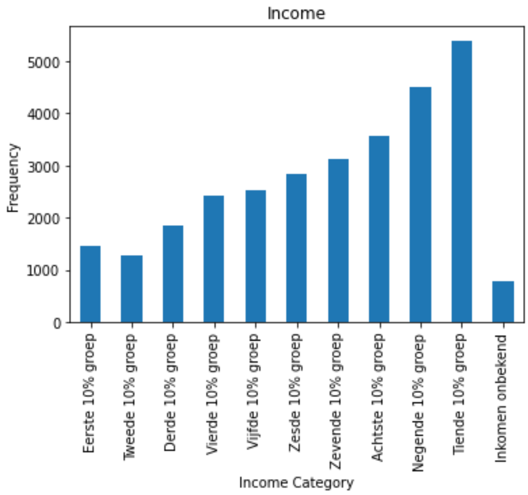


Figure D.4: Income class deciles

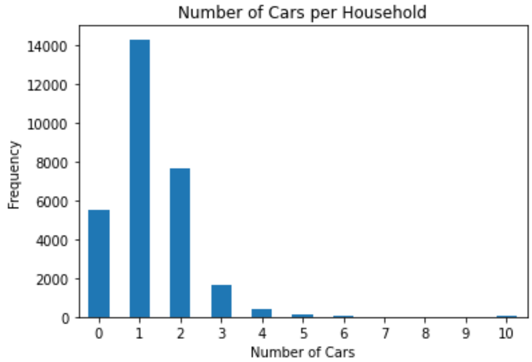


Figure D.5: Car ownership per household in the sample

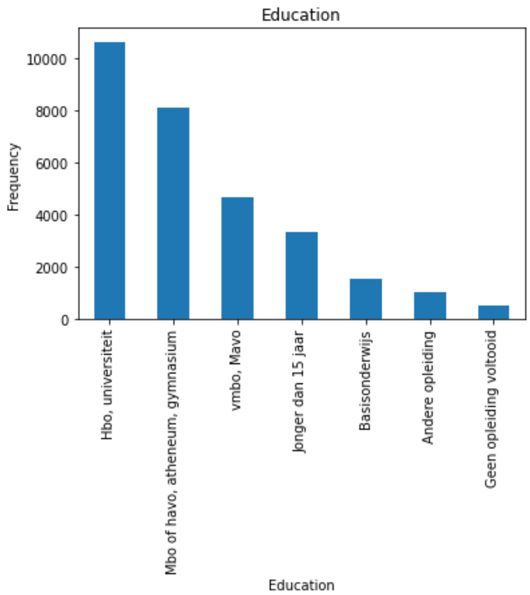


Figure D.6: Education level of respondents

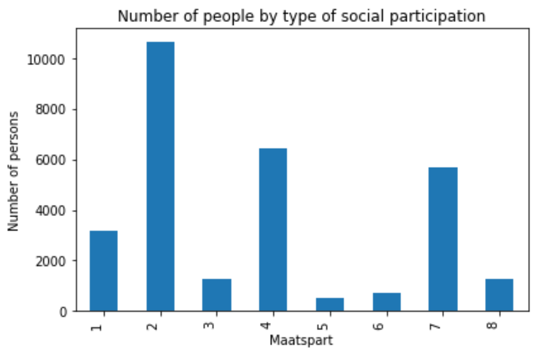


Figure D.7: Social participation of respondents. 1. Employed 12 to 30 hours per week, 2. Employed 30 hours or more per week, 3. Own household, 4. Student, 5. Unemployed, 6. Unfit for work, 7. Retired/VUT (Early retirement), 8. Other, 9. Unknown.

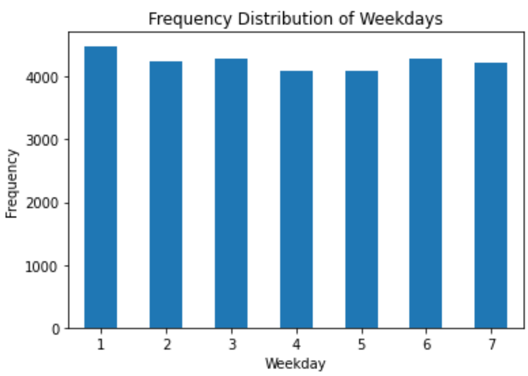


Figure D.8: Distribution of days the ODiN data is recorded.

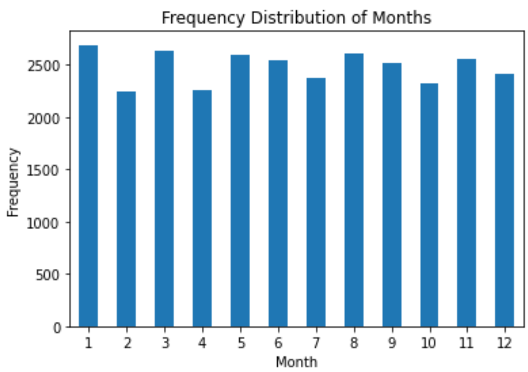


Figure D.9: Distribution of months the ODiN data is recorded.

E

Travel behaviour data

E.1. Data differences over the years

Table E.1 presents the differences in ODiN data across the years. The most notable distinction is between 2018–2019 and 2023, the post-COVID ODiN year. The 2023 data was included to ensure sufficient data, particularly for rural locations.

TB Data ZH	ZH 2018 + 2019 (n = 19840)	ZH 2023 (n = 9881)	Difference compared to 2018+2019 in ZH	NL 2019	NL 2023	Difference in NL compared to 2019
Car trips	1.25	1.12	-10,4%	1.26	1.21	-4%
PT trips	0.23	0.18	-22%	0.154	0.124	-19%
Active trips	1.32	1.36	3%	1.18	1.29	9%
Total trips	2.80	2.66	-5%	2.71	2.72	0%
Distance car	22.96	19.28	-16%	x	x	
Distance total	33.21	28.18	-15%	36.16	31.97	-12%

Table E.1: Differences between pre- and post-corona data

E.2. Difference in car ownership data

Table E.2 this table presents the differences due to the fact that lease-cars are not counted in the CBS data. It also gives the number of cars across ODiN data from respondents, which is not in line with the average measured by the KiM ”Atlas van de Auto” research (Martensen & Arendsen, 2024a).

BE cluster	KiM_onderzoek	CBS Data	ODiN Data
1	1.16	1	1.41
2	0.75	0.65	0.92
3	1.2	0.99	1.45
4	1.38	1.2	1.61
5	1.39	1.2	1.62
6	0.53	0.44	0.7
7	0.85	0.66	1.01
Mean	1.08	0.92	1.25

Table E.2: Differences between data sources, the KiM ”Atlas van de Auto” research, CBS Data and the ODiN dataset.

E.3. Data processing steps

Figure E.1 presents the method of the data processing, with the number of respondents on the left and the number of trips on the right. Due to the processing steps the number of ODiN respondents went from 41.856 to 29.721.

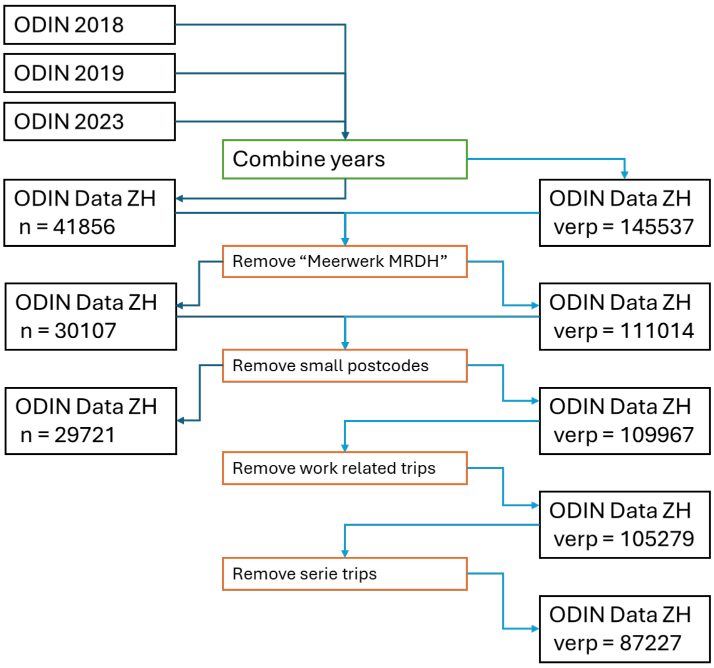


Figure E.1: Data processing steps, number of respondents (n) and number of trips (verpl).

E.4. Modal split home-work trips per distance

This subsection presents the modal split of all work related trip as cumulative per distance class of 2, 5 or 15 km. A potential walking/cycling, cycling/e-bike and e-bike distance. All data per trip motive could be found on this dataset webpage: maxkl.nl

Cluster 1	**Car**	**PT**	**Active**
2km	18% (n = 105)	0% (n = 1)	82% (n = 468)
5km	29% (n = 429)	2% (n = 32)	69% (n = 1010)
15km	43% (n = 1336)	8% (n = 237)	49% (n = 1536)
total	56% (n = 2977)	14% (n = 745)	30% (n = 1595)

Table E.3: Modal split home-work trips: Cluster 1

cluster 2	**Car**	**PT**	**Active**
2km	9% (n = 32)	2% (n = 8)	89% (n = 322)
5km	15% (n = 151)	6% (n = 64)	78% (n = 773)
15km	27% (n = 462)	14% (n = 229)	59% (n = 1005)
total	38% (n = 1151)	27% (n = 811)	35% (n = 1039)

Table E.4: Modal split home-work trips: Cluster 2

cluster 3	Car	PT	Active
2km	16% (n = 52)	0% (n = 0)	84% (n = 275)
5km	28% (n = 182)	0% (n = 3)	71% (n = 454)
15km	46% (n = 605)	7% (n = 90)	47% (n = 617)
total	59% (n = 1340)	13% (n = 291)	29% (n = 654)

Table E.5: Modal split home-work trips: Cluster 3

cluster 4	Car	PT	Active
2km	16% (n = 21)	0% (n = 0)	84% (n = 111)
5km	31% (n = 70)	3% (n = 7)	66% (n = 149)
15km	51% (n = 207)	2% (n = 7)	48% (n = 195)
total	70% (n = 630)	7% (n = 64)	23% (n = 210)

Table E.6: Modal split home-work trips: Cluster 4

cluster 5	Car	PT	Active
2km	19% (n = 26)	0% (n = 0)	81% (n = 110)
5km	27% (n = 68)	0% (n = 1)	72% (n = 180)
15km	49% (n = 307)	7% (n = 44)	44% (n = 281)
total	65% (n = 747)	8% (n = 90)	27% (n = 308)

Table E.7: Modal split home-work trips: Cluster 5

cluster 6	Car	PT	Active
2km	5% (n = 14)	1% (n = 3)	94% (n = 253)
5km	9% (n = 60)	15% (n = 101)	77% (n = 528)
15km	18% (n = 203)	23% (n = 258)	58% (n = 641)
total	24% (n = 465)	41% (n = 789)	34% (n = 650)

Table E.8: Modal split home-work trips: Cluster 6

cluster 7	Car	PT	Active
2km	13% (n = 19)	1% (n = 1)	86% (n = 122)
5km	20% (n = 90)	9% (n = 40)	71% (n = 315)
15km	27% (n = 240)	18% (n = 163)	54% (n = 481)
total	38% (n = 503)	25% (n = 333)	37% (n = 489)

Table E.9: Modal split home-work trips: Cluster 7

F

Validation cluster membership estimation

This appendix provides validation for the estimation of cluster membership. It presents the cluster membership probabilities of the respondents, allowing for a comparison with the estimated cluster membership probabilities.

F.1. Car ownership 0 and cluster membership

With a car ownership of 0 there are still individuals in a car-related cluster.

Cluster	Frequency	Percent	Cumulative Percent
Active only	2169	39,0	39,0
Car Only	389	7,0	46,0
Car + Active	282	5,1	51,1
No trip	1377	24,8	75,8
PT + Active	562	10,1	85,9
PT Only	675	12,1	98,0
Mix	109	2,0	100,0
Total	5563	100,0	

Figure F.1: Cluster membership of respondents who do not own a car.

F.2. Revealed and estimated cluster membership proportions.

Revealed cluster memberships with different variables. The estimations with the parameters could be done on website. Cluster 1: active, Cluster 2: car only, Cluster 3: car + active, Cluster 4: no trip, Cluster 5: train + active, Cluster 6: pt only, Cluster 7: mix.

F.2.1. BE indicators and TB cluster probability

Cluster modal

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	180	33,8	33,8	33,8
	2	87	16,4	16,4	50,2
	3	66	12,4	12,4	62,6
	4	80	15,0	15,0	77,6
	5	51	9,6	9,6	87,2
	6	52	9,8	9,8	97,0
	7	16	3,0	3,0	100,0
	Total	532	100,0	100,0	

Figure F.2: Filtered on: carown < 0.6 & carown > 0.4 & trainstationkm < 3 & trainstationkm > 1 & score < 250 & score > 230

Cluster modal

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	83	21,4	21,4	21,4
	2	162	41,9	41,9	63,3
	3	68	17,6	17,6	80,9
	4	51	13,2	13,2	94,1
	5	7	1,8	1,8	95,9
	6	10	2,6	2,6	98,4
	7	6	1,6	1,6	100,0
	Total	387	100,0	100,0	

Figure F.3: Filtered on: carown < 1.3 & carown > 1.1 & treinstationkm < 25 & treinstationkm > 10 & score < 60 & score > 40

Cluster modal

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	261	26,8	26,8	26,8
	2	319	32,8	32,8	59,6
	3	176	18,1	18,1	77,7
	4	141	14,5	14,5	92,2
	5	24	2,5	2,5	94,7
	6	36	3,7	3,7	98,4
	7	16	1,6	1,6	100,0
	Total	973	100,0	100,0	

Figure F.4: Filtered on: carown < 1.2 & carown > 1 & treinstationkm < 5 & treinstationkm > 2 & score < 110 & score > 90

F.2.2. Sociovariables TB and TB cluster probability

TBcluster

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,0	14	10,8	10,8	10,8
	2,0	66	50,8	50,8	61,5
	3,0	28	21,5	21,5	83,1
	4,0	12	9,2	9,2	92,3
	5,0	5	3,8	3,8	96,2
	6,0	2	1,5	1,5	97,7
	7,0	3	2,3	2,3	100,0
	Total	130	100,0	100,0	

Figure F.5: 30 <= Age & Age <= 40 & HHsize = 2 & HHAuto >= 2 & 6 <= Inkomen & Inkomen <= 8

TBcluster

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,0	23	11,6	11,6	11,6
	2,0	105	52,8	52,8	64,3
	3,0	47	23,6	23,6	87,9
	4,0	16	8,0	8,0	96,0
	5,0	2	1,0	1,0	97,0
	6,0	2	1,0	1,0	98,0
	7,0	4	2,0	2,0	100,0
	Total	199	100,0	100,0	

Figure F.6: 30 <= Age & Age <= 40 & HHsize = 2 & HHAuto >= 2 & 6 <= Inkomen & Inkomen <= 8

TBcluster

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,0	49	37,4	37,4	37,4
	2,0	10	7,6	7,6	45,0
	4,0	58	44,3	44,3	89,3
	5,0	3	2,3	2,3	91,6
	6,0	9	6,9	6,9	98,5
	7,0	2	1,5	1,5	100,0
	Total	131	100,0	100,0	

Figure F.7: 75 <= Age & Age <= 85 & HHsize = 2 & HHAuto = 0 & 4 <= Inkomen & Inkomen <= 6

TBcluster

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,0	146	28,6	28,6	28,6
	2,0	150	29,4	29,4	57,9
	3,0	104	20,4	20,4	78,3
	4,0	42	8,2	8,2	86,5
	5,0	23	4,5	4,5	91,0
	6,0	30	5,9	5,9	96,9
	7,0	16	3,1	3,1	100,0
	Total	511	100,0	100,0	

Figure F.8: 40 <= Age & Age <= 60 & HHsize = 2 & HHAuto = 1 & 8 <= Inkomen & Inkomen <= 10

TBcluster

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,0	131	25,8	25,8	25,8
	2,0	158	31,1	31,1	56,9
	3,0	91	17,9	17,9	74,8
	4,0	91	17,9	17,9	92,7
	5,0	11	2,2	2,2	94,9
	6,0	17	3,3	3,3	98,2
	7,0	9	1,8	1,8	100,0
	Total	508	100,0	100,0	

Figure F.9: 27 <= Age & Age <= 50 & 3 <= HHsize & HHsize <= 5 & HHAuto = 1 & 4 <= Inkomen & Inkomen <= 6