Technische Universität München

Physical Activity Assessment and Modelling using Household Travel Surveys

Thesis submitted in partial fulfilment of the degree
Master of Science in Transportation Systems

Advisers Qin Zhang, M.Sc.
Prof. Dr.-Ing. Rolf Moeckel
Professorship for Modelling Spatial Mobility
Department of Civil, Geo, and Environmental Engineering
Technical University of Munich

External adviser Dr. James Woodcock
Public Health Modelling
MRC Epidemiology Unit, School of Clinical Medicine
University of Cambridge

Author Corin Staves
corin.staves@tum.de

Submission date 14 October 2020

ACKNOWLEDGEMENTS

I would first like to express my gratitude to my supervisor, Qin Zhang, for her continued support and guidance, and for always being easily accessible to answer my queries. I would also like to thank Dr. Rolf Moeckel for joining our meetings and offering his ideas and guidance.

I would like to express my gratitude to my external supervisor, Dr. James Woodcock, for his continuous support, ideas, and guidance. Thank you to Anna Goodman for helping me with national travel survey data, and to Lambed Tatah for helping me understand the health impact estimation process. I would also like to thank the wider team at the MRC Epidemiology Unit for hosting my stay in Cambridge and for providing me with equipment, course materials in physical activity epidemiology, and partial funding through the GLASST project.

I am very grateful to Nadin Klomke and Dalma Alagha for supporting my stay abroad and offering partial funding through the Erasmus+ Traineeship scheme.

Finally, I would like to express my sincere gratitude to the German-American Fulbright-Kommission for initially supporting my studies at the Technical University of Munich, making this all possible.

SUMMARY

According to the World Health Organisation, about one third of the global population do not have sufficient levels of physical activity (PA). Active transport is viewed as an easy and effective way to introduce more PA into people's everyday lives.

In both transport and health science, there is increasing interest in assessing and modelling active transport and its associated health benefits. Household travel surveys (HTSs) could be a useful data source because they are more detailed than the PA questionnaires traditionally used in health studies. Transport modelling methods could offer new possibilities for forecasting the health impacts of active travel policy scenarios. However, these tools and techniques used by transport scientists might not be immediately suitable from a PA epidemiological perspective.

A literature review of health studies using HTSs and transport modelling tools revealed methodological uncertainties. For example, it was difficult to study PA during public transport trips. Compensatory behaviour (e.g. an increase in cycling causing a reduction in walking) is also poorly understood. Furthermore, HTSs and transport models generally only consider behaviour on a single day, whereas health guidelines and modelling tools generally consider week-long behaviour.

To better understand the potential for transport methods in health impact studies, this thesis investigates these uncertainties in detail. Specifically, it asks: **how can transport assessment and modelling methods be adapted to more accurately evaluate PA and its health impacts?** In the methodology, an adapted transport modelling framework is proposed. Indicators relevant to PA epidemiology were chosen; namely, the distribution of PA in the population and the estimated health impacts.

The first half of this thesis covers assessment. An exploratory approach compares three national HTSs to investigate how different data collection methods influence PA assessment. These were Germany's *Mobilität in Deutschland* and *Mobilitätspanel* (MOP) and England's *National Travel Survey*. Afterward, the MOP is explored in further detail to learn more about the distribution of PA in the population, determinants of PA, week-long behavioural patterns, and the effects of using a 1-day diary rather than a 7-day diary.

The second half of this thesis covers modelling, specifically trip generation and mode choice. First, the suitability of the trip-based paradigm is assessed. Afterward, using the proposed modelling framework, statistical analyses and discrete choice models are prepared from MOP data. Finally, to test the effectiveness of the proposed framework, predictions from the adapted mode choice model are compared with predictions from more traditional mode choice models.

The exploratory cross-dataset analysis revealed several possible HTS adaptations that would permit more accurate PA assessment. The proposed framework also proved to be useful for more accurately modelling PA and its health impacts. However, not all of the proposed adaptations in the framework could be proven to be beneficial with respect to the indicators specified. Nevertheless, the analyses provide insight into the complexities of active travel, such as mandatory-discretionary relationships, compensatory behaviour, and stability in travel decisions. These findings address some of the limitations brought up in previous transport health impact studies. They also offer policy potential by introducing new possibilities for defining active travel polices, targeting specific behavioural groups, and plausibly predicting policy impacts.

CONTENTS

Chapter 1. Introduction

1.1. Background

Transport systems play a major role in determining public health. Motor vehicle emissions are the main source of urban air pollution and thereby contribute to the 8.8 million air pollution related deaths each year by stroke, heart disease, lung cancer, and related conditions [1, 2]. Motor vehicle crashes cause approximately 1.4 million deaths worldwide each year and are the leading cause of death among people aged 5–29 [3]. Motor vehicle noise exposes many people to harmful levels and causes an estimated 50,000 heart attack deaths each year in Europe [4]. Furthermore, the transport sector is responsible for one quarter of greenhouse gas emissions in Europe and is therefore a major contributor to climate change and its associated long-term public health impacts [2, 5].

Sustainable transportation is becoming increasingly relevant to planners and policymakers worldwide, with high importance given to land use and transport systems that encourage public transport (PT) and active travel over private car. From a health perspective, such policies present a major opportunity. Not only can they reduce the health detriments of motorisation, but also offer health benefits through increased physical activity (PA).

Public health bodies around the world, including the World Health Organisation (WHO) and British and United States (US) governments, recommend that adults accumulate at least 150 minutes of moderate-vigorous physical activity (MVPA) per week [6–8]. Walking and cycling for transport qualify as being at least moderate intensity activity. Nevertheless, about 28% of adults worldwide (37% in high-income countries) do not achieve sufficient PA levels [9]. Insufficient PA increases the risk of cardiovascular diseases, cancer, diabetes, and other non-communicable diseases. Recent studies have estimated physical inactivity to be responsible for about 9% of premature deaths worldwide (5.3 million per year), making it one of the leading causes of global mortality [10].

Walking is the most popular form of PA [11]. Active travel, especially walking, is viewed by authorities as the easiest way to introduce PA to people's everyday life [7]. The more active travel somebody performs, the more beneficial it is for their health. However, this has been found to follow a law of diminishing returns; for example, the benefits associated with increasing walking from 0 to 20 minutes per week are greater than the benefits associated with increasing from 100 to 120 minutes per week [12]. Therefore, introducing new active travel to otherwise inactive individuals would be expected to offer the greatest population health benefits.

1.2. Defining Physical Activity

For PA epidemiologists, the most popular definition is “any bodily movements produced by skeletal muscles that result in energy expenditure” [13]. This definition differentiates it from the other forms of energy expenditure, namely resting energy expenditure (defined as “energy needed to maintain vital life functions during basal and sleeping conditions”) and thermic effect of food (defined as “energy required for purposes of digestion and the breakdown of food”) [13]. Physical activity energy expenditure (PAEE) is the most variable type of energy expenditure because it depends on individual decisions and behaviour rather than essential biological functions.

Absolute PAEE is the product of activity frequency, intensity, duration, and body weight. The intensity of an activity depends on the rate of oxygen uptake; in other words, how demanding it is. For example, walking is more intense then driving, and running is more intense than walking. The unit for absolute PAEE is energy divided by time (e.g. kilocalories per week). Note that with this measure, the same exact activity routine (e.g. walking 10 minutes per day) will generate a different PAEE for every individual, depending on body weight. Therefore, many PA studies exclude body weight from the calculation to arrive at a ‘volume’ of PAEE measure that depends only on activity frequency, intensity, and duration. This ‘PAEE volume’ measure is in units of energy divided by time and weight (e.g. kilocalories per kilogram per week).

PA is generally differentiated based on the setting, or domain, in which it is performed. Common domains are [13]:

- Occupational (e.g. manual-labour tasks, lifting objects)
- Household (e.g. yard work, childcare)
- Leisure (e.g. participation in sports, exercise)
- Transport (e.g. walking, cycling, navigating PT stations)

For a study to assess total PA, all domains would need to be captured. However, many studies focus on a single domain in detail to understand more precisely the determinants and health impacts of PA in that domain. For example, studies investigating strategies to encourage more walking and cycling focus on the transport domain.

1.3. Assessing Physical Activity

Standard assessment tools

Methods for assessing (i.e. measuring) PA vary widely in accuracy, precision and scope. The gold standard for assessing all-domain PA is doubly labelled water [14]. To perform this analysis, research subjects drink water enriched with ‘heavy’ isotopes of hydrogen and oxygen. Later, samples of urine, blood, or breast milk are taken to calculate the dissipation rate of these isotopes, which is then used to calculate total carbon dioxide production in the subject. This method can be highly accurate and reproducible, but it is expensive and requires stable conditions, making it infeasible for large-scale studies. Other objective PA assessment tools include indirect calorimetry, accelerometers, motion sensors, pedometers, and direct observation [13]. While these are more accurate than non-objective measures, the tools are still expensive, requiring stable conditions and consistent equipment.

Self-report questionnaires can be a low-cost and convenient alternative to these objective measures [15]. This makes them feasible for large-scale PA studies, including the global comparison by Guthold et. al. [9] mentioned previously. Global comparisons commonly use the international physical activity questionnaire (IPAQ) and global physical activity questionnaire (GPAQ). These aim to capture PA across all domains and offer unified global formats, making cross-cultural comparisons possible. The IPAQ asks respondents to consider specific activities performed in the past seven days, whereas the GPAQ asks about a ‘typical’ day or week. Some more detailed questionnaires use PA diaries in which respondents provide a “detailed or hour-by-hour or activity-by-activity record” [13]. However, in most questionnaires, respondents do not report each activity individually; instead, they answer questions about overall behaviour in a single sitting.

PA questionnaires are frequently tested against objective measures to assess their reliability and validity in different global contexts. A systematic review of 96 such studies by Helmerhorst et. al. [15] found that PA questionnaires are generally “acceptable” in terms of reliability, but their validity against objective measures is “moderate at best”. The American Heart Association noted that PA questionnaires are especially weak for investigating light to moderate intensity activities, such as walking [13].

Assessment using household travel surveys

Household travel surveys (HTSs) could be an alternative data source for assessing transport-domain PA. HTS respondents are required to complete travel diaries in which they document every occasion in which they interact with the transport system, including active travel. They document trips resulting from changes in location between activities (e.g. travelling from home to work) as well as simple roundtrip excursions that start and end at the same location (e.g. walking the dog, going for a bike ride). Trip mode, start/end times, speed, and distance are all captured. This information could be used to calculate transport-domain PAEE volumes for each respondent with less recall bias and greater precision than questionnaires like the IPAQ and GPAQ. Furthermore, HTSs are data-rich, often containing detailed information on travel purpose, general travel behaviour, attitudes, person-household interactions and sociodemographic attributes. This information could be useful for researchers to investigate and model the determinants of transport-domain PA. Using HTSs has the added advantage of reducing observer or social desirability bias because PA assessment is not their intended purpose.

Despite these potential advantages, there would be several pitfalls in using household travel surveys in this way. Obviously, their capacity to measure non-transport-domain PA is limited. It might be possible to estimate workplace-domain PA using the stated occupation (if available) and time spent working, or similarly leisure-domain physical activity using the detailed purpose description (if available) and activity time. However, the accuracy of these estimates would likely be poorer than the direct questions asked on PA questionnaires.

Another potential issue is that unlike the IPAQ and GPAQ, there are no global standards for household travel surveys [16, 17]. The definition of trip purposes and travel modes can vary from survey to survey. The methodology for administering the survey, calculating travel times and distances, coding roundtrips, and post-processing filtering also can also vary. These differences make cross-cultural comparisons difficult and possibly unreliable. It also means that a validation study performed on one HTS might not be applicable to another.

Finally, the time period of most HTS trip diaries might be too short for reliably assessing PA at the individual level. Large-scale HTSs are cross-sectional rather than longitudinal, so most travel diaries only record a single day [16]. However, a person's PA on a single day might not be representative of their overall behaviour. For example, consider somebody who cycles for 50 minutes three days per week. This person would exactly meet the 150-minute per week WHO recommendation. However, there is only a 3 in 7 chance of any of this cycling being captured in a single-day travel diary. If it is captured, then it would capture 50 minutes per day, which would translate to 350 minutes per week, a significant overestimation. In general, estimating PAEE using a single-day diary would be expected to underestimate the number of individuals with low (but not zero) PAEE while overestimating the number of individuals with zero PAEE or very high PAEE. This could affect population health estimates when drawing comparisons or testing scenarios. The extent of the error would depend on the variation (or conversely, the stability) in people's active travel behaviour, which could be investigated using a longitudinal survey.

1.4. Modelling Physical Activity

Transport models could help public health researchers and policymakers predict the impacts of active travel policy scenarios. Originating in the 1950s, transport models were traditionally oriented towards private car travel, with the aim of forecasting road congestion. However, the shift to sustainable transport has created increasing interest in modelling active travel [18]. There remains a concern that the traditionally car-oriented modelling methodologies might leave out information that, while not important for understanding car travel, could be crucial for understanding active travel.

Advancements in computational power have led to an onset of agent-based models, which apply stochastic discrete choice methods to simulate transport decisions at the individual level. Agent-based models are more flexible and allow researchers to use statistical models with a wider range of predictors, enabling more sophisticated behavioural approaches. Furthermore, the individual-level approach enables more detailed analysis of forecasts [19]. This could enable health researchers to better understand the distribution of active travel within different population segments and to estimate health impacts more precisely.

HTSs are a key data source for transport model development, especially trip generation and mode choice. Therefore, the pitfall associated with the lack of longitudinal travel survey data would also apply in modelling. If the model outputs a single-day forecast, the results would have been calibrated to a single-day travel diary. As a result, the model would not be able to capture day-to-day variation and the PAEE distribution would be distorted as previously described: low-PA individuals would be underestimated while zero-PA and high-PA individuals would be overestimated.

If the model outputs a multiple-day forecast (e.g. a week-long model) but continues to use traditional mode choice methods, the PAEE distribution could also be distorted. Mode choice models commonly apply the logit model. This type of model assumes that the error terms (i.e. the unobserved portion of utility) for each choice are independent and identically distributed (i.i.d.). This assumption can be unrealistic for long-term models because individual choices are actually correlated over time [20]. An example of this is presented for cycling. Typically, a substantial portion of cycling utility is unobserved and therefore part of the error term. This is because cycling strongly depends on individual preferences and tastes that are cannot be captured in travel surveys. Therefore, for a given individual, there could be large correlations between decisions in the error terms for cycling utility. An ordinary logit model would not be able to capture this because of the i.i.d. assumption. It cannot tell the difference between 1 person cycling 10 trips and 10 people cycling 1 trip. This could have significant implications for the PAEE distribution, likely causing low-PA individuals to be overestimated and zero-PA and high-PA individuals to be underestimated.

In summary, the potential pitfalls and consequences mentioned for single-day forecasts are opposite to those mentioned for multiple-day logit forecasts. Single-day forecasts cannot represent day-to-day *variability* in individual travel behaviour. Multiple-day forecasts using logit models cannot capture day-to-day *stability* in individual travel behaviour. This thesis hypothesises that both would need to be understood to accurately model active travel in a health context.

1.5. Research Question

This study investigates the potential for transport assessment and modelling methods that use HTSs. Specifically, it asks: **how can transport assessment and modelling methods using HTSs be adapted to more accurately evaluate transport-domain PA and its health impacts?** The central hypothesis is that traditional methods for assessing and modelling active transport from a travel demand perspective would not be suitable "as-is" from a PA epidemiology perspective, but the following adaptations could improve their accuracy:

- Capturing week-long behaviour instead of single-day behaviour
- Redefining trip purposes to be more relevant to PA
- Modelling the impact of mandatory travel on discretionary travel
- Restricting the modes available to each individual
- Incorporating more complex substitution patterns between modes

Before testing these adaptations, an exploratory analysis will investigate how different HTS data collection and assessment methodologies can influence how PA is captured in transport. The aim of this is to gain further insight into potential inaccuracies caused by differing data collection practices.

The primary indicator of this study will be the distribution of transport-domain PA volume in the population. Because of the nonlinearity of PA health impacts, both the mean and the shape of the distribution are important, so both will be considered. A secondary indicator will be the health impacts. These indicators will be elaborated on further in the methodology chapter.

1.6. Overview of this Thesis

Chapter 2 provides a review of academic studies have that used HTSs to assess and model active travel from an epidemiological perspective and discusses core uncertainties in their methodology. Chapter 3 describes the chosen data sources, explains the methodology for estimating PA volume and health impacts, and proposes a transport modelling framework that aims to model PA more effectively.

Chapters 4 and 5 cover assessment. Chapter 4 explores and compares three national HTSs to understand how their different approaches impact the ability to capture transport-domain PA. Two HTSs capture the same target population, Germany, enabling direct comparison between their assessment techniques. A third HTS captures the population of England, which will offer insight into the ability of HTSs to be used for cross-cultural PA comparisons. One of the HTSs offers a full longitudinal 7-day travel diary. Chapter 5 explores this longitudinal HTS in detail to calculate the full distribution of weekly transport-domain PA and investigate how mandatory travel patterns influence discretionary travel and PA. This chapter also investigates how the PA distribution and health impact estimates would be different if PA were assessed using a 1-day rather than a 7-day diary.

Chapters 6 and 7 cover modelling, with trip generation in chapter 6 and mode choice in chapter 7. These chapters apply the adapted model structure introduced previously in chapter 3. Statistical modelling methods are applied to develop trip generation and mode choice models using an agent-based approach. Particular focus is given to defining trip purposes, modelling the impact of mandatory on discretionary travel, modelling substitution patterns, and restricting the modes available to each individual. Chapter 7 will evaluate the effectiveness of the adaptations to the model structure by comparing mode choice forecasts with more conventional modelling approaches. Many other aspects of travel demand modelling are also relevant to PA; however, this thesis only investigates trip generation and mode choice because these modelling components are generally the most reliant on HTS data.

Chapter 2. Literature Review

A literature review was conducted to investigate how previous studies have applied HTS data to investigate transport-domain PA. Section 2.1 reviews assessment studies, while section 2.2 reviews modelling studies. Section 2.3 describes their key methodological uncertainties while introducing relevant studies and tools from the transport field. Finally, section 2.4 describes the research gaps that will be addressed in this thesis.

2.1. Assessment Studies

A large number of studies globally have used HTSs as a data source to learn about active travel patterns. However, these have mostly been from a travel demand perspective rather than a PA epidemiological perspective [21].

The first study identified to have used an HTS from a PA epidemiologic perspective was conducted by Merom et. al. [22] in 2010. The study pointed out several advantages of HTSs over traditional PA questionnaires, including the greater accuracy associated with travel diaries, the ability to differentiate cycling from other forms of moderate-vigorous PA, and the ability to differentiate active travel by purpose (e.g. utilitarian vs. recreational). In their methodology, they used data from a single-day Sydney HTS to identify whether respondents met various health-enhancing thresholds for walking and cycling.

In 2011, Pucher et. al. [23] applied the methodology from Merom. et. al. to the US 2001 and 2009 national HTSs to assess changes in transport-domain PA over time. These US surveys used a single-day travel diary but also provided some insight into longitudinal travel behaviour by including two short recall questions on the number of walking and cycling trips made over the previous week.

Buehler et. al. [24] applied the same methodology again to create an international comparison on transport-domain PA in the United States and Germany. Their data sources were the same US surveys from 2001 and 2009 and the German *Mobilität in Deutschland* surveys from 2002 and 2008. The authors noted that while household travel survey methodologies vary widely, they were fortunate in that the US and Germany surveys were highly comparable. However, they did point out differences in the way roundtrips were coded (1 trip in Germany but 2 in the US) and noted that only US respondents were specifically prompted not to neglect short walks. The authors believed that this methodological difference may have led to an underestimation in the difference in transport-domain PA between the two countries.

In 2015, Fishman et. al. [21] used the single-day Dutch HTS to understand the contribution of transport-domain PA to meeting PA guideline requirements and the impacts of weather, urban structure, and socio-demographic factors. Unlike the previous three studies, this study multiplied the duration of walking and cycling by intensity to calculate a PAEE volume for each individual. This acknowledged that cycling was a higher intensity – and therefore more beneficial – activity than walking. The authors acknowledged that while this made it possible to calculate PAEE volumes for each individual, they could not calculate total PAEE as this would require information such as body weight which the travel survey did not provide.

2.2. Modelling Studies

Aggregate modelling

Since 2005, the WHO regional office for Europe has been developing its health economic Impact assessment (HEAT) tool [25]. This tool provides a methodology for modelling health impacts of active transport scenarios, including PA. It was designed to be accessible to non-health experts, thereby helping policymakers and transport professionals to assess the health impacts of their scenarios. Several studies have applied HEAT using travel surveys as their main data source. Rojas-Rueda et. al. [26] used HEAT to estimate the health impacts of shifting Barcelona car trips to cycling. Grabow et. al. [27] used HEAT to estimate the health benefits of replacing short car trips with active modes in the US Midwest. In the UK, an adaptation of the HEAT tool is included in the Government's official transport appraisal guidance, offering a straightforward methodology for UK transport planners to assess health impacts of their proposals [28].

HEAT only provides estimates for all-cause mortality. In contrast, a 2009 study by Woodcock et. al. [29] estimated disease-specific health impacts of PA. The study used data from London and United Kingdom (UK) travel surveys to investigate health impacts of imagined transport forecasts for 2030. For example, one scenario involved increasing walking and cycling levels in London to be similar to cities like Copenhagen or Amsterdam. The tools from this study later developed into an integrated transport and health impacts model (ITHIM). ITHIM has been used for other transport health impact studies based on scenarios in California ([30]), San Francisco (Maizlish et. al [31]), Nashville (Whitfield et. al. [32]), Los Angeles ([33]), England and Wales (Woodcock et. al. [34]), and São Paulo (Sá et. al. [35]).

Several studies have used similar methodologies to estimate active-travel health impacts without using ITHIM or HEAT. Hartog et. al. [36] used their own methodology to assess the health trade-offs of replacing short car trips with cycling. Xia et al. [37] used an older WHO framework to estimate the health impacts of a mode shift to active travel in Adelaide. Tainio et. al. [38] used a methodology based on previous studies to estimate active travel health trade-offs between PA and air pollution exposure.

A systematic review of health impact studies and their results was prepared in 2015 by Mueller et. al. [39]. This review synthesises results from the studies mentioned here. It also includes results from other health impact studies that were not included here as they did not primarily rely on travel survey data (for example, Rojas-Rueda et. al. [40] and Woodcock et. al. [41] used cycle hire data to investigate the health impacts of cycle hire programmes in Barcelona and London respectively).

Most transport health impact studies do not define scenarios based on modelled transport forecasts, but rather they assume direct changes to HTS data (e.g. '50% of car trips under 1.5km switch to walking') or aggregate changes (e.g. 'the miles walked per capita is doubled'). However, some studies mentioned that modelled data would have been more desirable as it could have provided a better scope [29, 41] or greater scenario flexibility [27].

Microsimulation modelling

More recent studies have estimated health impacts of transport scenarios that were modelled to some extent using microsimulation. However, this literature review could not identify any studies that used a full travel demand model (e.g. a four-step model) to inform health impact scenarios.

Ahmad et. al. [42] used logit models to estimate how a denser urban structure in Nashville would increase the likelihood of choosing to walk and cycle. They then calculated health impacts of this 'high-density' scenario using ITHIM.

Lovelace et. al [43] developed the propensity to cycle tool (PCT), a planning support system for England and Wales to identify which commuting routes have the highest potential for cycling uptake. The PCT inputs home-work commute origin-destination (OD) pairs from the UK census. It then uses logit models to determine which OD pairs are most likely to be cycled. Different scenarios apply different models (for example, a 'go Dutch' scenario uses a logit model estimated using a Dutch HTS). Then, a cycle-specific route choice model is applied to show which routes have the highest cycling potential. Finally, health impacts (including PA) are estimated using an approach similar to HEAT [44].

Woodcock et. al [45] developed the impacts of cycling tool (ICT). This is like the PCT, but it uses UK HTS data instead of census data. By using HTS data, it can capture trips across all purposes instead of just commuting. In their methodology, trips and distances from the travel diary are taken as input, and then mode choice is estimated using logit models. As in the PCT, different mode choice model structures are applied for different scenarios. For each scenario, health impacts associated with increased PA were estimated using a methodology based on a recent systematic review.

All three of these studies applied an individual-level microsimulation approach. Their authors consider microsimulation to be advantageous because of its greater realism, enhanced scenario capabilities, ability to model different population sub-groups, and ability to calculate health impacts more precisely.

2.3. Key Methodological Uncertainties

Public Transport Access and Egress

Most of the reviewed studies stated the importance of including PA from PT access and egress. However, access and egress data from HTSs can be inconsistent and difficult to capture [22]. As a result, some studies could not capture PA from PT trips and mentioned this as a limitation [27, 31].

Those that did include PT active travel did so in different ways depending on the data available. Some studies made a blanket assumption that all PT trips included a specific (constant) walking distance [26, 40]. Other studies imputed access and egress distances based on supplementary questions at the trip or household level [23, 24]. Other studies had access to stage-level survey data [21, 45]. Other studies had precise OD coordinates available, which they could use in combination with GIS to determine the distance to PT stations [22, 30]. The methods for calculating PT active travel, and the validity of these methods, were not described in detail (if at all).

A recent study by Chaix et. al. [46] used objective measures to capture walking during PT trips. Their study combined a travel diary with data from GPS sensors and accelerometers. Based on their algorithm, they found that PT access and egress accounted for about 37% of all walking in urban areas, and about 30% of all walking in suburban areas. They concluded that PT access and egress is indeed a key component of transport-domain PA that should not be neglected in health studies.

Compensatory Behaviour

Between domains of physical activity

HTSs capture transport-domain PA but provide little to no information about PA in other domains (workplace, leisure). Therefore, many health impact studies supplement HTS data with PA questionnaire data to estimate the distribution of other-domain PAEE. For aggregate modelling studies, the distribution is generally differentiated by age and sex and then added to age- and sex-specific transport-domain PAEE distributions. For the agent-based ICT, individual records in the travel survey were randomly matched to similar individuals in a PA questionnaire [45]. However, matching is generally only possible by age and sex.

It may be the case that a person who is more active in the transport PA domain is compensating by being less active in other PA domains. Unfortunately, this relationship is difficult to capture, so transport health impact studies assume other-domain PA is constant across scenarios. This has caused many authors to express concern about 'compensatory balancing' between different domains of PA, and how it may cause an overestimation of the health benefits of active travel [29–31, 36, 40].

There is little evidence on this substitution phenomenon due to a lack of data and understanding [36, 39, 47]. The systematic review by Mueller et. al. [39] found that some studies showed significant health benefits for increased active travel "even after adjusting for other domains of PA", but other studies showed "uncertainty as to how much active travel contributes to total PA". They found that recent evidence from longitudinal studies showed little substitution between active travel and other PA domains, so they concluded that the assumption of constant other-domain PA is plausible.

De Hartog et. al. [36] performed a brief sensitivity analysis into the effects of this type of compensatory behaviour on their estimated health impacts of switching car trips to cycling. They found that the health benefits of switching to cycling would only be outweighed by the costs if at least 89% of the additional cycling PA was due to compensatory balancing.

Within the transport domain

Studies in the transport field have found that individual travel time budgets "tend to remain constant over time" [48, 49]. Therefore, given that active travel is slower than non-active travel, individuals who switch to active modes may compensate for this lost time by choosing closer destinations or taking fewer trips. De Hartog et. al. [36] pointed out that switching from driving to cycling would cause people to choose destinations closer to home. However, this was not included in the modelling, so it was stated as a limitation. Woodcock et. al. [45] considered the opposite effect for those switching from walking to cycling, who would experience time savings and may therefore take longer or more frequent trips. This would increase PAEE but it was not included in their modelling, so it was stated as a limitation.

Merom et. al.'s [22] descriptive analysis coded trips by purposes separately and concluded that recreation and commuting purposes were the main drivers of cycling uptake. However, none of the reviewed modelling studies separated trips by their purpose, making substitution between mandatory and discretionary trips irrelevant. One study (Lovelace et. al. [43]) only considered home-based work and education trips. The authors claimed this might have caused an underestimation of health benefits because their scenarios would have probably increased cycling in discretionary trips as well.

Between modes

Substitution between modes is also relevant to health impacts. For example, a policy that increases cycling would likely draw new cyclists from all other modes, including walking. However, walking involves more PAEE than cycling for a given distance [43]. This means switching from walking to cycling would disbenefit health, *ceteris paribus*. Lovelace et. al. [43] and Woodcock et. al. [45] accounted for this in their cycling studies by incorporating the reduced PAEE for those who switched from walking. They assumed equal substitution, meaning all modes are equally likely to switch to cycling. In his cycle-sharing study, Woodcock et. al. [41] used data from surveyed cycle-hire customers to determine which modes were being substituted. They found that most new cyclists came from PT (47%) or walking (31%). In another cycle-sharing study, Rojas-Rueda et. al. [40] tested two different substitution possibilities for the new cyclists (0% previously walked vs. 30% previously walked) and found the difference in health impacts to cyclists to be minor (10.5 vs. 12.5 deaths avoided). Other studies compared several different substitution possibilities [26, 30, 34]. However, some studies assumed that 100% of new cyclists came from driving [27, 36].

Piatkowski et. al. [50] were critical of the assumed substitution patterns in health impact studies (particularly those of Grabow et. al. [27]) as they lack an empirical foundation and may therefore create misleading results on the health impacts of walking or cycling. The authors performed an intercept survey on walkers and cyclists in five US cities and found that substitution between active modes and car trips “ranged between 25% and 86%”. This substitution varied significantly based on location, age, car use, and helmet use. The authors also said that trip purpose is likely a significant factor in the substitution rate, but they did not test this.

In transport modelling, the ordinary logit mode choice model assumes equal substitution [51]. This is because it exhibits a property of *independence from irrelevant alternatives* (IIA). A popular way to relax this IIA property in mode choice modelling is to use nested logit models, which group similar alternatives into nests [20]. The nesting parameter of each nest describes the extent to which the unexplained portions of utility are correlated. For example, in a study on the determinants of active travel modes, Ton et. al. [52] built a nested logit model based on the Dutch mobility panel survey. They could not find any significant nesting parameter between walking and cycling and concluded that they should not be nested together. Moeckel et. al. [53] developed a nested logit mode choice model using German HTS and arrived at the same conclusion. However, Kattak et. al. [54] developed a nested logit mode choice model for Pittsburgh and found a suitable nesting structure in which walking, cycling, and car driving were in the same nest.

Relative Risks and Dose-response Relationships

Once PAEE volumes are calculated for each population subgroup (or individual if agent-based), health impacts are estimated by means of dose-response relationships. These are functions relating exposure with relative risk (RR). In this case, exposure is a PAEE volume. RR is defined as the ratio of the probability of a particular health outcome occurring in an exposed group versus a non-exposed group [55]. The health outcome could be a particular disease (e.g. diabetes, dementia, heart disease) or it could be all-cause mortality (death). For example, an all-cause mortality RR of 0.75 would indicate that a person would be 0.75 times as likely to die within a given time period if exposed (i.e. a heath benefit). Modelling studies ideally require a continuous dose-response function, with RR decreasing gradually as PAEE volume increases.

Dose-response relationships can be identified using cohort studies. These are studies that follow respondents over extended periods of time to understand how their PA patterns impact health. However, cohort studies usually only offer point (i.e. discrete) RR estimates associated with specific levels of exposure [29]. In 2000 Andersen et. al. [56] published results from 2 cohort studies from Copenhagen on the health impacts of cycling. This source had been heavily cited in older health impact studies, including previous versions of HEAT [25, 26, 36, 40].

Recent systematic reviews have combined results from many cohort studies to try to identify a continuous dose-response function [12, 29, 57]. There is a general consensus that RR has the largest decrease at smaller PAEE levels but flattens at larger PAEE levels (rule of diminishing returns). This means the health benefits of increasing somebody's walking by 20 minutes per week will be greater for somebody with baseline walking of 0 minutes per week than for somebody with baseline walking of 100 minutes per week. However, the evidence for the precise shape is conflicting, especially towards the higher-PAEE end of the curve.

A 2011 systematic review by Woodcock et. al. [57] found that some studies suggested a linear relationship (equal benefit at all levels), while other studies suggested there was no benefit at higher PAEE levels. Based on this review of 22 cohort studies, the authors estimated a curve using power transformations and found that a 0.25 power transformation had the best fit.

Unless a linear curve is used, PAEE at baseline can significantly influence health impact estimates. If baseline PAEE volume is underestimated, then the slope of the RR curve will be overestimated. This could lead to an overestimate of the change in RR under scenarios, and therefore an overestimate of health benefits. However, estimating baseline PAEE requires estimating PAEE across all domains. This is problematic because of difficulty matching transport and other-domain PAEE data as described previously. Because of this, linear functions are the most popular choice for health impact studies because their constant slope makes baseline PA irrelevant, eliminating the need to estimate non-transport-domain PAEE [39, 58]. However, it is acknowledged that this linear curve is probably unrealistic.

More recent systematic reviews of cohort studies have aimed to identify curvilinear dose-response functions for transport-domain PAEE independently of other-domain PAEE. The systematic review by Woodcock et. al. [57] investigated walking-specific dose-response relationships and found that a 0.375 power curve had the best fit. Later, Kelly et. al. [12] investigated walking and cycling specific dose-response relationships. Using data from 21 cohort studies, they built and compared linear, log-linear and 0.25, 0.5, 0.375, and 0.75 power curves. These are shown in figure 1. The results from this review are especially advantageous because the functions do not depend on other-domain PAEE. These have been applied in several transport health impact tools and studies, including the most recent version of HEAT [38, 42, 43, 58].

The systematic review by Mueller et. al. [39] noted a large variation in the type of function chosen in different health impact studies, with the most common being linear. Given the uncertainties regarding the shape of the dose-response curve, some studies have performed sensitivity analyses by testing different functions. Woodcock et. el. [29] tested linear, linear with cutoff, and square root exposure transformations, noting that the shape of the dose-response relationship can be "especially important" in the health impact estimations. De Hartog et. al. [36] also tested different dose-response functions and also found a significant impact. Rojas-rueda et. al. [26] compared HEAT's linear function with Woodcock et. al.'s [57] curvilinear functions and also found a significant difference (62.8 vs. 38.5 deaths avoided).

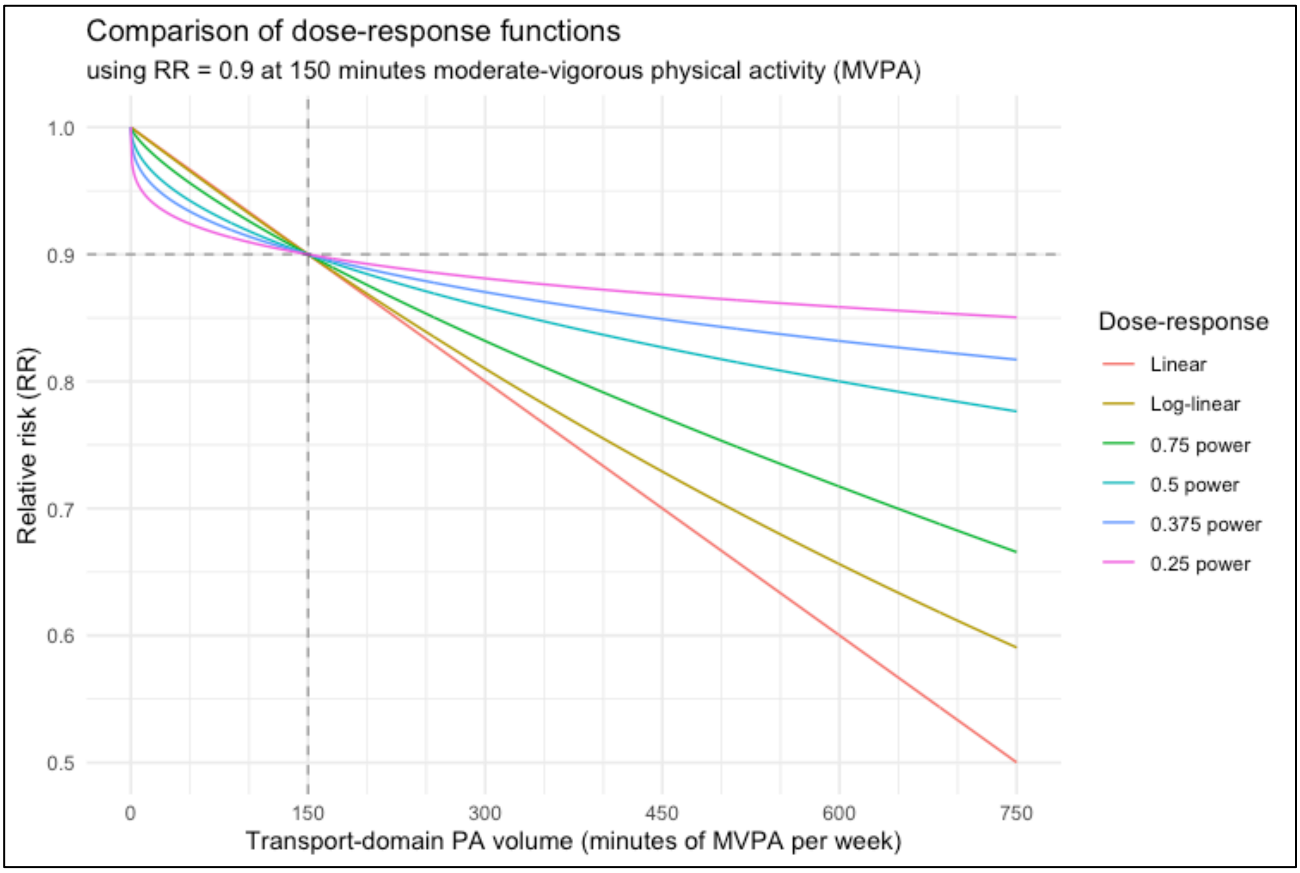


Figure 1: Comparison of dose-response relationships (source: adapted from Kelly. et. al. [12]**)**

Length of the Travel Diary

Public health guideline recommendations use a week-long timescale. PA questionnaires are usually designed to capture PA over the course of a full week (as in IPAQ) or a *typical* day or week (as in GPAQ). Both these types of questionnaires were used in the cohort studies investigated by Woodcock et. al. [57] and Kelly et. al. [12]. Health impact tools such as ITHIM, HEAT, PCT, and ICT use week-long timescales to be consistent with these PA requirements, questionnaires, and cohort studies.

On the other hand, travel diaries usually only cover a *specific* day [16]. As discussed in the introduction, the distribution of PAEE volumes across in population could be significantly different for a single day than it would be for a 'typical' day or a week. Woodcock et. al. [29] said that using a 1-day diary would likely overestimate the variation in physical activity. Merom et. al. [22] acknowledged that the lack of longitudinal data was a limitation of their study. They explained that this made it impossible to assess the regularity of travel behaviour, which is an "important component of health". Pucher et. al. [23] compared single-day travel diary results to individual-level questions about active trips made in the previous week. In their aggregate-level comparisons of socioeconomic subgroups, they found similar results using both the daily and weekly data. They concluded that using single-day data was reasonable for population health estimates but not for individual-level estimates as the “randomly selected travel day may have been atypical”.

Most of the reviewed health impact studies relied on single-day travel diary data [27, 29, 30, 32, 33, 35, 42]. In order to be compatible with the week-long health impact tools, assumptions were made about the shape of the week-long distribution. Woodcock et. al. [29], who used a 1-day London diary, calculated a coefficient of variation using data from a 7-day nationwide diary and assumed this would be the same for the London population. However, the 7-day data was only available for longer walking trips, so an assumption was made that the variation in long trips was the same as for short trips. Sá et. al. [35] used a similar approach but, due to a lack of data, had to apply a coefficient of variation from a high-income setting to a low-income setting. They pointed this out as a limitation of their study. Maizlish et. al. [31], who used a 2-day diary, tested alternative coefficients of variation in their sensitivity analysis. They found that the alternative coefficients only caused a 2% difference to health impacts in their scenario.

As discussed in the introduction, individual-level variation can significantly impact the PAEE distributions using modelled data. Lovelace et. al.’s PCT [43] relies on census data, which only includes commuting OD zone pairs and the “usual mode of travel to work”. This presents the same issue as using single-day diary data. It is not possible to capture how many work trips were cycled, so the authors apply an average number of cycling trips for those whose ‘usual mode’ is cycling and zero cycling trips to everybody else. This means the PCT cannot capture variation in active commuting behaviour: Individuals who cycled 40% and drove 60% of the time would be neglected entirely as cycling is not their ‘usual mode’; similarly, individuals who cycle an above-average number of commutes could not be captured properly as an average is applied to everybody.

Woodcock et. al.’s ICT [45] uses logit models to estimate mode choice over a 7-day period. As discussed in the introduction, this could also distort the modelled PAEE distribution because logit models cannot capture *stability* in individual behaviour. The authors address this issue by modelling mode choice over 2 steps. In this case, the models are binary as they are only concerned with whether or not the mode is cycling.

- **Step 1:** Individual-level logit model to estimate who would consider cycling during their 7-day diary. In this step, individuals are designated as either “cyclist” or “non-cyclist”.
- **Step 2:** Trip-level logit model to estimate which trips would be cycled (restricted to those individuals who were designated as “cyclists” in step 1).

By differentiating between cyclists and non-cyclists in the trip-level model, it prevents the number of individuals who cycle from being overestimated. Therefore, the number of individuals with zero cycling PAEE is not underestimated, overcoming a key limitation of using logit models for mode choice over longer periods. Nevertheless, for those who are cyclists, this structure would still not be able to differentiate between regular cyclists (e.g. cycles every day) and irregular cyclists (e.g. cycles once per week).

Relevant studies in the Transport Field

In the transport field, extended panel surveys have been used to investigate the variability in individual travel behaviour. Heinen and Chatterjee [59] provide a literature review of many such studies. Another relevant study was conducted by Crawford [60], who used a British 7-day travel diary to investigate relationships between dominant and secondary commute modes. She found that 83% of individuals do not vary their commute mode, but walkers and cyclists were more likely to vary their commute mode than car drivers. Beltman [61] used data from a 6-week Dutch travel diary to calculate a mode variation index (MIX) for each individual, differentiating by trip characteristics and socioeconomic characteristics. He found that MIX increased as trip length increased, and he also found that MIX was largest for leisure trips and lowest for commute trips.

In mode choice modelling, a powerful tool for incorporating the stability (and/or variability) of individual choices over time is the *mixed logit* model [20, 62]. This is a generalisation of the ordinary logit model in which random coefficients are introduced to the utility function. These coefficients can be defined with a great deal of flexibility; for example, they could be defined as varying between individuals but constant for all choices of a given individual. This would enable correlations across choices of the same individual, overcoming this key limitation of ordinary logit models. Cherchi and Cirillo [63, 64] and Cherchi et. al. [65] built mixed logit models from the German and Swiss 6-week *Mobidrive* travel diaries. In all studies, incorporating these random coefficients substantially improved their model results over ordinary logit. Another key finding of these studies was that there is significantly more interpersonal variation across days of the week than across weeks, suggesting that a single-week survey could be a reasonable estimate of overall travel behaviour. Thomas et. al. [66] used a 4-week Dutch travel diary to construct mixed logit models to investigate interpersonal variation and its relationship with time, space, trip purpose, and dominant mode. They found that intrapersonal variation was largest for short trips (< 2km) and for recreation trips. They also found that differentiating by dominant mode significantly improved model fit.

Models covering extended periods for transport forecasting studies are rare. This literature review identified only four studies that developed such models. Beltman [61] used an ordinary logit model based on a 6-week travel diary to forecast mode choice over the 6-week period. He found that the model substantially overestimates variation in individual mode choices over the 6-week period. This is not surprising because of the i.i.d. assumption of ordinary logit models meaning they cannot capture stability in individual behaviour (as discussed previously). Mallig and Vortisch [67] estimated a mode choice model using a 7-day German travel diary as part of the *mobiTopp* 7-day travel demand model. They also used ordinary logit, meaning the overestimation in variability would also be present (although this was not tested). Woodcock et. al. [45] defined a cycling model that partially overcame this issue by applying a 2-step model described in the previous section. Vij [68] was the only study identified to have applied a mixed logit model for mode choice forecasting. As part of his PhD, he estimated a mixed logit model using discrete random coefficients (i.e. a *latent class* model) using 6-week *Mobidrive* data. He then used the model to identify different modality types and test policy scenarios.

This literature review found that mixed logit models are much more complex than ordinary logit models and often involve difficulties with identifying a suitable structure, training, convergence, and interpretation. It is therefore not surprising that only one forecasting study was found to have applied one. Written guidance typically recommends using a simpler model structures if at all possible [20, 62].

2.4. Research Contribution

This literature review includes important findings that help to answer the research question. In contrast to PA questionnaires, which often follow global standards, HTSs use different methods to collect travel data, which could impact PA assessment. This makes it difficult to use them for cross-cultural PA comparisons. Therefore, most transport-domain PA assessment and modelling studies using HTSs have focused on a particular region, each requiring a substantially adapted methodology depending on the data available. There has been little attention to the suitability of the chosen dataset with regard to PA assessment, or the suitability of applying the same methods in another context or for cross-cultural comparisons.

There is a consensus that closer integration with transport models, especially microsimulation models, would be beneficial for health impact studies. However, the few studies that used microsimulation travel-demand modelling approaches had heavily simplified and adapted methods, based on single-purpose binary mode choice models only (e.g. cycle vs. other). No studies were identified that differentiated between purposes or modelled multiple modes, as in traditional travel demand modelling.

There is also a consensus that substitution between different types of PA could significantly influence health impact estimates. However, there are major research gaps in this area. Unfortunately, relationships *between* different PA domains cannot be investigated effectively using HTSs alone. However, substitution patterns *within* the transport domain can be. Many studies discussed substitution between work and leisure in *total* PA as a limitation of their work. Only two studies discussed the possibility of a relationship between work and leisure PA *within the transport domain* (i.e. mandatory and discretionary trip purposes), but they did not investigate the relationship further. Substitution between active modes and non-active modes has also been discussed in several studies; however, modelling studies have so far assumed either complete substitution between a non-active and active mode, or equal substitution between all modes, both of which might overestimate health benefits.

Finally, there is little consensus on which dose-response relationship is most suitable for estimating PA health impacts. However, there a consensus that the type of curve chosen can substantially influence health impact results.

This thesis builds from the reviewed literature to address the described research gaps in the following ways: First, this study will explore how different types of datasets (and corresponding PA assessment methodologies) influence the assessment of transport-domain PA. Second, this study will use descriptive and statistical approaches to investigate the potential of multi-modal mode choice and multi-purpose trip generation approaches for modelling PA. Third, this study will investigate the relationship between mandatory and discretionary travel, the potential for compensatory behaviour, and its implications for PA. Finally, this study will investigate substitution between active modes, and the possibility of overestimating health impacts if it is ignored. Throughout this study, specific attention is given to the impact of using different types of dose-response functions. Focusing on these research gaps will guide the analysis of the adaptations proposed through answering the research question to investigate their potential for future research.

Chapter 3. Data Sources & Methodology

3.1. Data Sources

This study uses data from three recent nationwide HTSs conducted in Germany and England. These are all self-report surveys in which households complete travel diaries designed to capture every occasion in which a household member changes location (or otherwise interact with the transport system, for example to walk a dog). All three surveys are part of large-scale government-maintained studies, and their datasets are national standards for transport analysis and modelling building by researchers and practitioners in their respective countries. A summary table of each dataset and relevant attributes for this study is given in table 1.

Deutsches Mobilitätspanel (Germany)

The Deutsches Mobilitätspanel (MOP) is carried out by Karlsruher Institute für Technologie, on behalf of the *Bundesministerium für Verkehr und digitale Infrastruktur* [69]. The MOP is a panel survey that has been repeated every year since 1994. MOP recruits 650 new households per year and aims to keep households in the panel for three years to achieve about 1850 households in each yearly cohort. In the MOP, individuals aged 10+ within each household complete a 7-day travel diary. Of the three datasets considered for this study, the MOP is the only one that captures all trips (including short walks) over a seven-day period. However, the MOP is limited in size, scope, and precision of measurements.

The MOP is the primary data source for analysis and modelling in this study. Only data from 2016 onward was included because of a significant change to the data collection methodology that occurred between 2015 and 2016. Furthermore, as the year-to-year panel nature of the MOP was not relevant for this study, records in each year were assumed to be independent from the previous year.

Mobilität in Deutschland (Germany)

Mobilität in Deutschland (MiD) is carried out by the infas *Institut für angewandte Sozialwissenschaft* (institute for applied science) on behalf of the *Bundesministerium für Verkehr und digitale Infrastruktur* (Federal Ministry of Transport and Digital Infrastructure) [70]. It is a large-scale cross-sectional travel survey originating in 1976 that is carried out approximately once every eight years. The most recent survey, MiD 2017, was carried out between 2016 and 2017; 156,420 households across Germany were interviewed. The advantages of MiD 2017 are in its size, scope, and precision of measurements and data. However, MiD respondents only complete a single-day travel diary.

Being a single-day survey, it is not possible to use the MiD travel diary to investigate week-long PA at the individual level. However, it is useful for aggregate analysis, especially because of the high quality of the data.

National Travel Survey (England)

The National Travel Survey (NTS) is carried out by NatCen Social Research on behalf of the UK Department for Transport [71]. It is a large-scale travel survey originating in 1965, and since 1988 it has become a continuous survey that is carried out every year. On average about 8500 households are interviewed each year. While some households may carry over from year to year, their household IDs are not preserved in the data so each year's records must be assumed to be independent. Like MiD, the NTS is strong in its size, scope, and precision of measurements. Like MOP, a seven-day travel diary is completed. However, respondents are only required to record walks of less than 1.6km on one of the seven days.

Only NTS data from 2014 onward was used for this study because of a significant change in the data collection methodology that occurred in 2013. With NTS data it is not possible to conduct individual-level analysis or modelling when short walks are involved. Therefore, the primary purposes of using NTS data in this study is to provide an international comparison for the German data at the aggregate level. Its limitations are also used as a framework for understanding the limitations of major travel surveys in the context of physical activity. For example, to better understand the short walk limitation, an analysis into the relevance of short walks in physical activity is conducted in part 2.

Table 1: Overview of household travel surveys used for this study

	Mobilität in Deutschland	**Deutsches Mobilitätspanel**	**National Travel Survey**
Country	Germany	Germany	England
Year	2016–2017	2016–2018	2014–2017
Households	156,420	5,452	29,179
Individuals	316,361	11,225	70,038
Trips	960,619	207,550	912,002
Ages with trip data	All ages	Age 10+ only	All ages
Travel diary period	1 day	7 days	1 day (walks ≤1.6km) 7 days (all other trips)
Stage level data available	Rarely (only 4% of individuals)	Never	Always
PT Access Distance Available	Home end only	Home end, Education end, Work end	Home end only

3.2. Selected Attributes and Filtering

Unlike PA questionnaires, there are no global standards for travel surveys. All three datasets relied on different methodologies and data collection procedures. The data from the three surveys were harmonised as much as possible to enable a meaningful comparison in chapter 4. This section explains that harmonisation process.

Chapters 5, 6, and 7 also use the MOP attributes that are described in this section.

Included Attributes

The relevant variables from each dataset are given in table 2. Unit conversions were applied were necessary (e.g. miles to kilometres, walk-minutes to kilometres, or day-km to week-km). A walking speed of 4.8 kilometres per hour was assumed based on the guidance in HEAT [58].

Table 2: Variables used for cross-dataset comparison

Variable	Mobilität in Deutschland (MiD)	Deutsches Mobilitätspanel (MOP)	National Travel Survey (NTS)
HOUSEHOLD ATTRIBUTES			
Nearest bus stop	*[bus28]*	*[HALTBUSZ]*	*[WalkBusTRACC]*
Nearest tram or metro stop	*[tram28]*	Minimum of *[HALTSTRZ] and [HALTUZ]*	*[StationKmTRACC]*
Nearest regional rail station	*[bahn28]*	Minimum of *[HALTSZ] and [HALTZUGZ]*	*[StationKmTRACC]*
Nearest intercity rail station	*[bahn28]*	*[HALTZUGZ]*	*[StationKmTRACC]*
PERSON ATTRIBUTES			
Age	*[HP_ALTER]*	Difference between *[JAHR]* and *[GEBJAHR]*	*[Age_B01ID]*
Sex	*[HP_SEX]*	*[SEX]*	*[Sex_B01ID]*
Distance between work/study place and nearest public transport station	---	*[FUSSMIN]*	---
TRIP ATTRIBUTES			
Trip Origin	**1st Trip:** *[W_SO1]* **Remaining Trips:** Destination of previous trip	**1st Trip:** assume 'home' unless respondent selected 'URLAUB' on that day **Remaining Trips:** Destination of previous trip	*[TripPurpFrom_B01ID]*
Trip Destination	*[W_ZWECK]*	*[ZWECK]*	*[TripPurpTo_B01ID]*
Trip Distance (km)	*[wegkm_imp]*	*[KM]*	*[JD]*
Trip Main Mode	*[hvm_diff2]*	*[VMDIW]*	*[MainMode_B03ID]*
STAGE ATTRIBUTES			
Stage Distance	*[ET_KM]*	---	*[SD]*
Stage Mode	*[ET_VM]*	---	*[StageMode_B03ID]*

Harmonising Trip Modes

The options for trip modes were different for each dataset. Table 3 shows which how the main modes from each dataset were categorised in the unified dataset.

Table 3: Assigning uniform trip modes

Assigned Main Mode	Mobilität in Deutschland (MiD)	Deutsches Mobilitätspanel (MOP)	National Travel Survey (NTS)
Car Driver (carD)	• Moped / mofa • Motorcycle (driver) • Car (driver)	• Moped / motorbike • Car (driver)	• Household car – driver • Non-household car – driver • Household motorcycle – driver • Non-household motorcycle – driver
Car passenger (carP)	• Motorcycle (passenger) • Car (passenger)	• Car (passenger)	• Household car – passenger • Non-household car – passenger • Household motorcycle – passenger • Non-household motorcycle – passenger
Public Transport (PT)	• City / regional bus • Tram • U-bahn / Stadtbahn / suspension railway • S-bahn / regional train • Long-distance train • Long-distance bus (regularly scheduled) • Long-distance bus (chartered)	• City / regional bus • Long-distance bus / coach • Tramway / light rail / metro • Urban rail / regional train • Long-distance train	• London stage bus • Other stage bus • Public express bus / coach • Excursion / tour bus • London underground • Surface rail • Light rail • Other public transport
Cycle	• Normal bicycle • E-bike / pedelec	• Normal bicycle • E-bike / pedelec	• Bicycle
Walk	• On foot	• On foot	• Walk, less than 1 mile • Walk, 1 mile or more
Other	• Car-sharing vehicle • Taxi • Anrufsammeltaxi or on-demand bus • Lorry (driver) • Lorry (passenger) • Ship / ferry • Aeroplane • Other	• Aeroplane • Ship • Lorry • Horse-drawn carriage • Motor home • Inline skates, kickboard • Other	• Private (hire) bus • Household van / lorry – driver • Non-household van / lorry – passenger • Household van / lorry – passenger • Non-household van / lorry – passenger • Other private transport • Air • Taxi • Minicab

Harmonising Trip Purposes

- Trip purposes were assigned using the trip's origin and destination. As the possible choices varied between datasets, purposes did not match exactly. The assessment part of this study differentiates between the following purposes: work, education, accompanying, shopping, recreation, recreational roundtrip, and other. For the modelling part, a more detailed investigation into trip purpose coding is given in section 6.1.

Table 4 shows how the purposes in each dataset assign were assigned to one of these. With the exception of roundtrips, trip purposes were also differentiated according to whether they were:

- Home-based (HB): Trips starting *or* ending at home (all purposes)
- Non home-based (NHB): Trips starting *and* ending away from home (work / other only)

Table 4: Assigning uniform trip purposes

Assigned purpose	Mobilität in Deutschland (MiD)	Deutsches Mobilitätspanel (MOP)	National Travel Survey (NTS)
Work	• Reaching the workplace • Official / business • Business trip	• Trip to work • Work related trip	• Work • In course of work
Education	• Reaching the training centre / school • School or pre-school • KITA / Kindergarten • Lessons (non-school)	• Trip to kindergarten, school, university	• Education
Accompanying	• Dropping off / picking up / accompanying others • Accompanying adults	• Picking someone up, dropping someone off	• Escort home • Escort work • Escort in course of work • Escort education • Escort shopping / personal business • Other escort
Shopping	• Shopping	• Shopping, running errands	• Food shopping • Non-food shopping
Recreation	• Leisure / free time • Sports activity / sports club • Visiting / meeting friends	• Leisure, hobby	• Eat/drink with friends • Visit friends • Other social • Entertain/public activity • Sport: participate
Recreational roundtrip	[NOT CAPTURED SEPERATELY]	• Trip starting and ending at the same location, e.g. round trip, stroll, walk, jogging tour, bike tour	• Day trip • Just walk
Other	• Private errands • Other	• Other • Trip back to hotel, second home • Trip to second home • Other private activity	• Personal business medical • Personal business eat/drink • Personal busines other • Holiday: base • Other non-escort

Filtering

The following persons and their trips were eliminated from the harmonised dataset:

- Persons aged under 10. This was to ensure consistency across datasets, since only individuals age 10+ completed trip diaries in the MOP.
- Persons with ‘unknown’ mobility (i.e. it is unknown whether any trips are taken)
- Persons with ‘unknown’ age
- Persons with ‘unknown’ sex

For those individuals that were included, all of their trips were considered for the cross-dataset analysis. This includes trips that are generally disregarded in trip-based modelling procedures (e.g. trips terminating at home, trips with unknown purposes, trips with implausible speeds). This was to make sure all possible sources of PA were captured at the assessment stage.

Additional Variables for Modelling

For the trip generation and mode choice modelling using MOP in chapters 6 and 7, the following additional MOP variables were used:

Household attributes

- Total size: [*HHGRO*]
- Number of children: sum of persons under 16 in *P*, *POT*, and *KIND* files
- Urban or rural: [*IDREGIOSTAR2*]
- Number of cars: [*PKWHH*]

Person attributes

- Driver’s licence: [*FSPKW*]
- Owns bicycle: [*FAHRRAD*]

Additional filtering was performed to avoid model bias due to irregular or unknown behaviour. Therefore, in addition to the filters described in the assessment stage, the following individuals and their trips were eliminated:

- Any individual with at least one ‘unknown’ trip purpose
- Any individual with at least one ‘unknown’ trip mode
- Any individual who is on holiday during any part of the travel diary. These were identified using the [*URLAUBxx*] variables in the person file.

3.3. Sources of Physical Activity

Physical Activity Outside the Transport Domain

This study does not use other-domain PA. This is consistent with more recent studies that have considered transport-domain PA independently of other-domain PA, such as Kelly et. al. [12], Lovelace et. al. [43], and Ahmad et. al. [42]. As discussed in the literature review, this has implications for the dose-response relationship, which will be discussed later in section 3.5.

Physical Activity in Trips with Main Mode Walking or Cycling

All walking and cycling main mode trips were considered as sources of PA.

Physical Activity in Trips with Other Main Modes

Costa et. al. [72] found that the PA intensity of travelling as a car driver, car passenger, or PT rider are significantly less than moderate-intensity and therefore do not contribute to the 150 minutes of moderate-intensity PA recommended by public health bodies. Therefore, only PA from walking and cycling was considered.

However, many trips are multi-modal. For example, PT journeys usually involve a walking component on either end. The literature review concluded that to understand health it is desirable to include walking and cycling from other-main-mode trips, especially PT trips. However, the availability and quality of this information varied greatly depending on the dataset. To gain an understanding of these differences, this study attempted to capture as much information as possible from each dataset regarding other-main-mode walking and cycling. The following sections describe how other-main-mode PA was captured in the datasets.

Trips with stage-level data

The British NTS requires respondents to report multi-modal trips in stages, each with a corresponding distance and mode. Stage reporting was also possible in the German MiD survey, but only 4% of respondents chose to do it.

Where this stage-level data was available, the walking and cycle stages of other-main-mode trips were included as sources of PA.

For many PT trips with stage-level data, respondents failed to report the stage covering the distance between their home and their home-end PT stop. If this home-end distance was known, a walk stage was assumed and also included as a source of PA.

Trips without stage-level data

If stage-level data wasn't available, other information in the datasets made it possible to predict the PT access modes and distances. All three surveys allowed respondents to select multiple modes for each trip. All three datasets also included information about distances from households to the closest PT stop or station. This made it possible to estimate whether an individual walked or cycled to their home-end PT stop, and the distance they would have covered. The estimation method utilized a rule-based approach based on the combination of modes selected for each trip.

The German MOP also included distances between individuals' work or education end PT stop and their place of work or study. Therefore, it was possible to estimate walk and/or cycle travel on this end as well.

3.4. Quantifying Physical Activity

As discussed in the introduction, PAEE volume is the product of frequency, duration, and intensity.

Frequency

Official PA guidelines, as well as ITHIM and HEAT, consider weekly volumes. Therefore, trips and stages that were only recorded on a single day in the travel diaries were multiplied by a factor of 7. This was the case for short walks in the NTS, and all trips in MiD.

The impacts of using such a factor to estimate the week-long PA distribution will be investigated in section 5.3.

Duration

Trip durations are available in HTSs. However, the methodology for determining travel times can vary between studies, leading to possible survey-specific reporting biases [30]. To avoid this possible source of error, durations are estimated using reported distances only. This is consistent with the methodology of many reviewed health impact studies.

Reported trip distances are converted to durations using average speeds. For more precise health impact estimates, speeds could be differentiated by age and sex (e.g. in [29, 31, 34]). However, for this study It was desirable not to do this as it could confound comparisons across different sexes and age groups. Therefore, a single average speed was chosen for each mode. This was 4.8 km/h for walking (from guidance in HEAT [58]) and 13.9 km/h for cycling (from a cycle speed investigation in [44]).

Intensity

PA intensity is usually described in units of metabolic equivalent tasks (METs). A single MET is defined as an oxygen uptake of 3.5 millilitres per kilogram of body weight per minute, which is approximately the rate of energy expenditure while sitting at rest [8]. Higher intensity activities have higher MET values. It is also common to use marginal METs (mMETs), defined as total METs minus 1 [45]. Using this definition, mMETs can be thought of as total energy expenditure *above* rest. The WHO defines moderate intensity PA to be 3–6 METs, or 2–5 mMETs.

In reality, the intensity of walking and cycling varies based on age, sex, geography (e.g. hilliness), path condition, equipment condition (e.g. of shoes or bicycle), health, body type, frequency of breaks, speed, acceleration, baggage, riding/walking style, and many other factors. While most of these attributes are not available in HTSs, some transport health impact studies can differentiate METs by age and sex (e.g. in [34]). To avoid confounding comparisons, this study uses a single MET value for each mode. This was 5.44 mMETs for cycling and 3.61 mMETs for walking, which were the median intensities obtained from a recent objective study on commuters by Costa et. al. [72].

PA Volume

Multiplying intensity by duration yields a total volume of PA, commonly with units 'MET-hours' or 'MET-minutes'. Multiplying this by a frequency would yield a PA volume per unit time, with units like 'MET-hours per week' or 'mMET-hours per week'.

This study uses mMETs instead of METs, which makes it easier to visualise and interpret data because only activities above rest are included. For example, a person performing an entirely sedentary activity would acquire 0 mMET-hours. This study only captures transport-domain PA, so only walking and cycling are figured into the final PAEE volume.

Public health bodies recommend a minimum of 150 minutes of moderate PA per week. Kelly et. al [12] and Woodcock et. al. [45] assume moderate-intensity PA to be about 4.5 METs (or 3.5 mMETs), which is the centre of the 3–6 MET range defined as moderate PA by the WHO. Multiplying this value by 150 minutes, the WHO minimum PA guideline translates to 11.25 MET-hours per week, or 8.75 mMET-hours per week. For additional health benefits, the WHO recommends at least 300 minutes of moderate intensity PA [8], which would translate to 22.5 MET-hours per week, or 17.5 mMET-hours per week.

The distribution of transport-domain PA volume in the population, in mMET-hours per week, is the primary indicator for this study.

3.5. Estimating Health Impacts

A key aim of this study is to understand how health impact estimates might be affected when the distribution of transport-domain PA is incorrect. Recall from the literature review that there is consensus of non-linearity in the dose-response relationship for PA. Because of this non-linearity, changing the shape of the PA distribution could have significant effects on the health prediction even if the mean value is constant.

Recall from section 2.3 that RR is the ratio of a health outcome occurring in an exposed group versus a non-exposed group. In this case, the exposure is PA volume (mMET-hours per week), and the outcome is all-cause mortality (i.e. death). Dose-response relationships were estimated using the all-cause mortality RR curves from Kelly et. al. [12]. These curves consider transport-domain PA independently of other-domain PA, eliminating the need to include non-transport-domain PA in the calculation. For consistency, this study assumes RR = 0.9 at 8.75 mMETs for all curves, like in figure 1. Applying these functions converts the PA volume into an RR value for each individual.

Expected health outcomes differ drastically by age and sex. For example, the risk of death is much higher for 80-year-olds than for 20-year-olds. Because of this, most health impact studies separate the population into age- and sex-specific groups before making further health calculations. This shifts the focus to older individuals, for whom changes have the largest impact. However, the goal of this study is to investigate PA for the population as a whole without prioritising particular groups. Therefore, to avoid confounding the results due to age- and sex-specific differences in the data, this study investigates health impacts for the population as a whole without separating into age and sex-specific groups. Age and sex are still included as predictors of travel behaviour.

To estimate health impacts of scenarios, the RRs of the scenario are compared to the RRs at baseline. Specifically, a population attributable fraction (PAF) can be calculated using the equation

$$PAF = \frac{\sum_{i=1}^{n} P_i RR_i - \sum_{i=1}^{n} P_i RR'_i}{\sum_{i=1}^{n} P_i RR_i}$$

where P_i is the (statistical) weight of each individual record i, RR_i is the baseline RR for each individual i, and RR'_i is the scenario RR for each individual i (equation adapted from [73]) . Put in other words, the **PAF is the percent difference in weighted mean RR between the baseline and scenario.** For all-cause mortality, the PAF can simply be multiplied by the mortality rate in the population group (e.g. "deaths per year") to determine the total health benefits (e.g. "deaths avoided per year").

This is clearly a heavily simplified approach for estimating the health impacts of transport-domain PA. Many of the reviewed health impact studies estimate health impacts in more detail by using age and sex specific outcomes and by considering specific health outcomes rather than all-cause mortality. However, the goal here is not to capture health impacts as precisely as possible. For this study, the most interesting variable in the health impact estimation process is the shape of the dose-response function that converts the PA volume distribution to an RR distribution. This is because different functions could cause the inaccuracies in the PA distribution to have different effects on mean RR and PAF. Therefore, all of the possible dose-response functions by Kelly et. al. are tested and compared.

These health impact estimators, specifically mean RR and PAF, are the secondary indicators for this study.

3.6. Proposed Modelling Framework

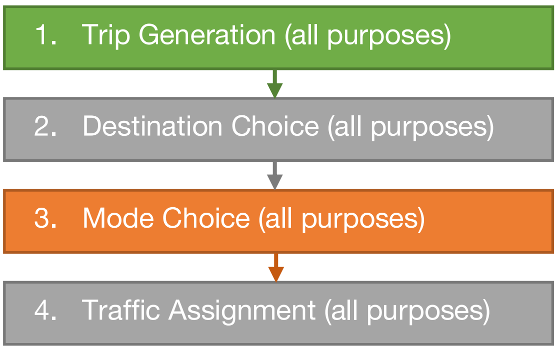


Figure 2: Traditional 4-step model structure

The traditional four-step model structure is shown in figure 2. Based on the research question and findings from the literature review, an adapted modelling framework is proposed, shown in figure 3. This section describes the motivation behind the modifications.

First, this structure proposes modelling travel behaviour over seven days, instead of just a single day or peak hour. As discussed in the literature review, longer period models have been constructed before by Beltman [61] (6 weeks), Mallig and Vortisch [67] (7 days), Woodcock et. al. [45] (7 days), and Vij [68] (6 weeks). Modelling a 7-day period will make the transport model consistent with health impact data and models. It will also ensure that the full variation in individual transport-domain PA over the week is captured.

Second, this structure introduces a primary commute model early in the process. This is defined as the mode used most frequently for commutes over the course of the week (same as in Lovelace et. al. [43]). It is similar to the definitions used by Crawford [60] and Thomas et. al. [66], except in this case it encompasses strictly work and education trips. Primary commute mode is estimated at the individual level. Trip-level estimations for all purposes remain at the mode choice stage, with primary commute mode being an independent variable. It is hypothesised that estimating primary commute mode before the main mode choice can introduce more stability into individuals' active travel decisions.

Third, this structure proposes estimating travel behaviour for mandatory trips before discretionary trips. Therefore, the number of mandatory trips and primary commute mode become independent variables for discretionary trip generation and mode choice. This will make it possible estimate how changes to mandatory travel behaviour would influence discretionary travel behaviour.

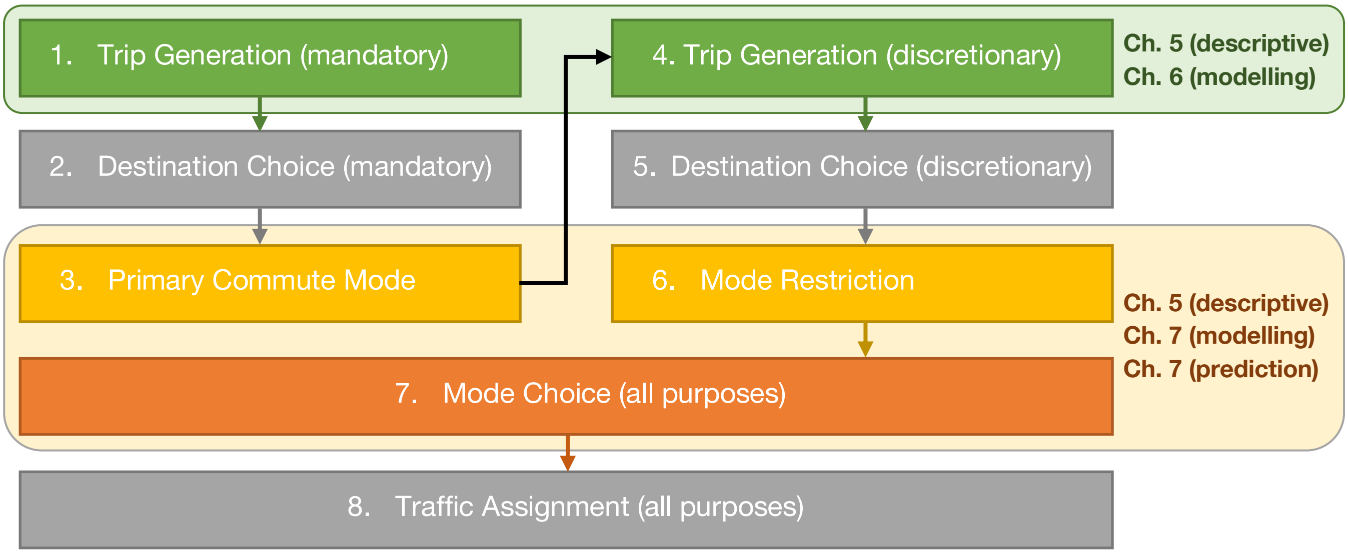


Figure 3: Modified 4-step model structure

Finally, this structure proposes introducing a mode restriction model before main mode choice. This model restricts the choice set for each individual, ensuring that the proportion of individuals using particular modes (e.g. walkers, cyclists, drivers) – and therefore the proportion of active individuals – is realistic and calibratable. The mode restriction model can be thought of as answering the question: “*which mode(s) would the agent be willing to use during the travel week?”* It was designed based on to the two-step approach of the 7-day cycling model by Woodcock et. al. [45], which differentiates cyclists from non-cyclists (described in section 2.3). However, unlike that model, which is a binary mode restriction, this model will consider multiple choice sets covering all potential modes.

This study only investigates trip generation and mode choice because these are the modelling components most reliant on HTS data. Chapter 5 investigates the impacts of mandatory behaviour (trip generation and mode choice) on discretionary behaviour using a descriptive approach. Chapter 6 uses statistical methods to estimate the trip generation model components independently of mode choice. Similarly, chapter 7 uses statistical methods to estimate and predict mode choice independently of trip generation. The simplified model structures for independently investigating trip generation and mode choice are shown in chapters 6 and 7 respectively. These simplified structures are based on full modelling framework in figure 3, but a full model structure is not estimated or run in this thesis.

All modelling and plotting for this study was performed using the R statistical computing language [74]. Plots were generated using the *ggplot2* package [75] unless otherwise specified.

Chapter 4. Aggregate-level Assessment using Three Datasets

This chapter synthesises data from all three household travel surveys to explore patterns in PA and PA-relevant travel behaviour. The goal of this analysis is to inform how different datasets affect the assessment of transport-domain PA, and to identify which data collection methodologies are most suitable for transport-domain PA assessment. The results from this chapter will also help to identify key relationships in transport-domain PA (e.g. age, sex, trip purpose) that will help to inform and support modelling decisions in later chapters.

Because of differing time scales, only aggregate-level analysis is possible for this comparison. This means average values can be calculated (e.g. 'total trips / total persons', or 'cycling mMET-hours per week / total females') but the shape of the distribution cannot be calculated.

4.1. Travel Behaviour Overview

Modal shares from each dataset are shown in figure 4. The mean number of trips per person per week, by mode, is shown in figure 5. Average distances by each mode are shown in figure 6.

To validate the data processing methodology, the modal splits from figure 4 and mean trips and distances from figures 5 and 6 were compared to results from official MiD and NTS publications [70, 76]. The findings aligned closely (within 3 units). However, because of the way variables were filtered and unified (described in section 3.2), it was not expected for these values to align exactly.

The modal share (figure 4) is similar for Germany and the UK. Notable exceptions are car passenger (about 1.4x higher in the UK), walking (about 1.2x higher in the UK), and cycling (about 5.9x higher in Germany). Total trip rates and distances across all modes in the UK were significantly smaller than in Germany. This discrepancy is not intuitive and could indicate poorer diary participation in the UK rather than a difference in travel behaviour. Nevertheless, these results match the values from publications in their respective countries.

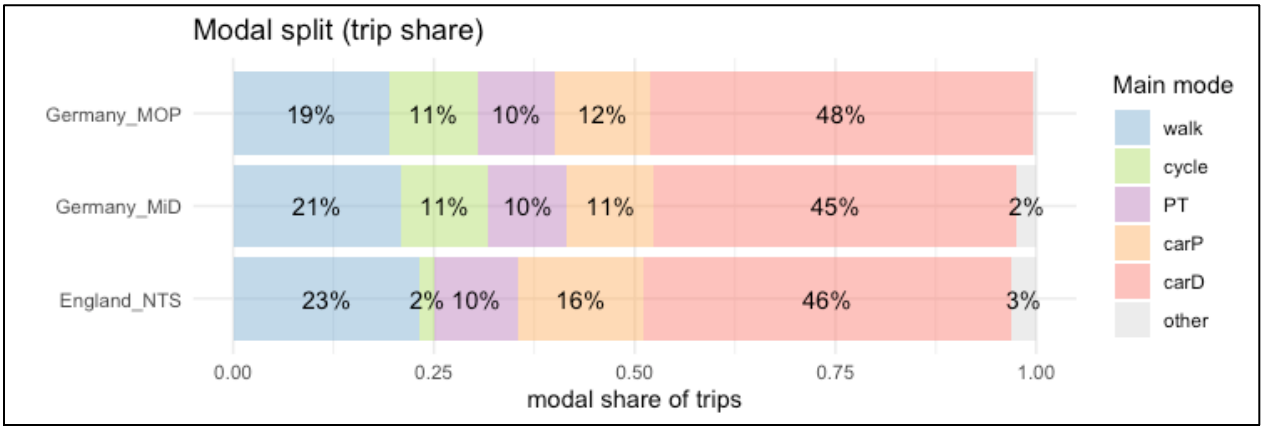


Figure 4: Modal split (trip share)

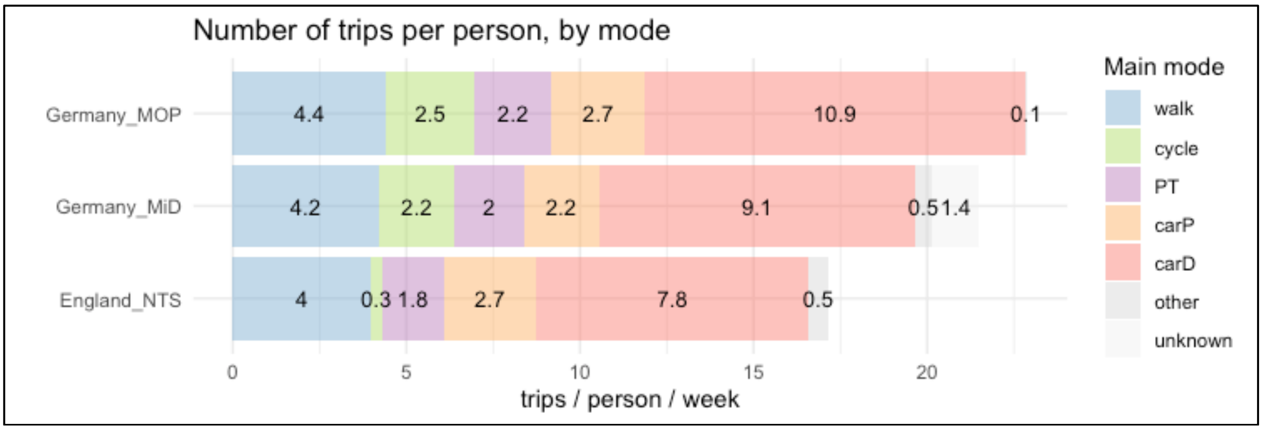


Figure 5: Number of trips per person, by mode

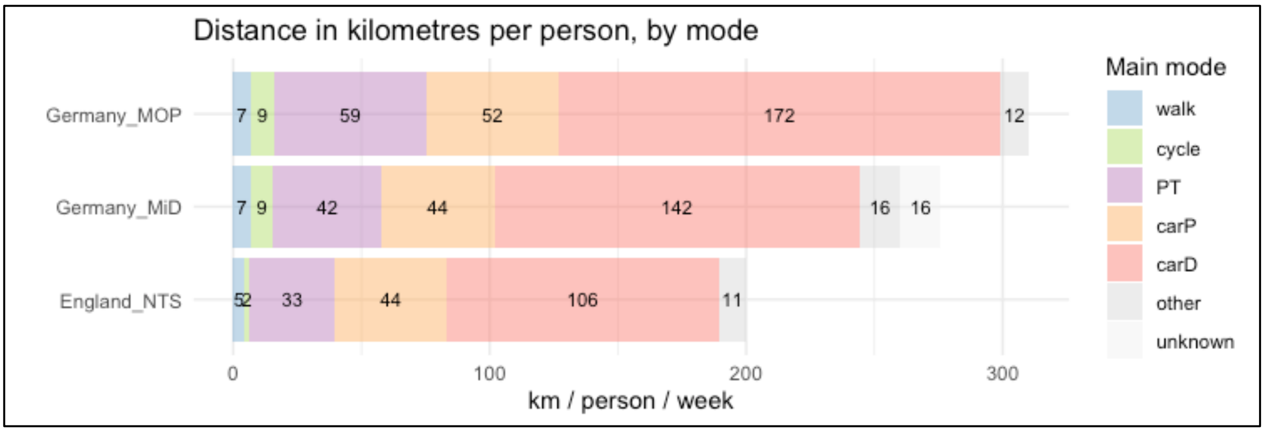


Figure 6: Distance in kilometres per person, by mode

The MOP and MiD yielded similar results for relative mode share. This is expected as these two surveys intend to capture the behaviour of the same population (i.e. German residents). However, the absolute distance and number of trips was substantially higher in the MOP than MiD. There are a number of potential reasons for this:

- The MOP did not differentiate between individuals who took no trips over the survey period and individuals who took trips but didn't complete the trip diary. Persons who recorded no trips were excluded from the MOP dataset entirely. However, in the MiD and NTS individuals were able to record that they were non-mobile.
- MiD respondents were only able to report the first 8 or 12 trips in their single-day diary, whereas MOP respondents could report up to 24 trips per day.
- Different surveying methodologies and levels of participation of smaller-scale panel respondents versus large-scale survey respondents.

Despite this discrepancy, the total distance walked and cycled per person (most important for PA) were about the same in both the MOP and MiD.

Figures 7 and 8 present a further differentiation by trip purpose. Figure 7 shows the mode share of trips (similar to figure 4), and figure 8 shows total distances (similar to figure 6). These figures reveal that car is used for the majority of trips for all purposes except education and recreation.

In both the UK and Germany, car is the dominant mode for all purposes except education and recreation. Active modes are more popular in education, shopping, and recreation than other purposes. Where recreational roundtrips are considered separately, an important observation can be made. The total distance travelled for recreational roundtrips is relatively small compared to other purposes; however, this purpose is dominated by active modes. This is crucial because round trips are often disregarded in transport models. If the goal is accurately capture to transport-domain PA, it would seem important not to disregard them.

NHB work trips account for many more person-km in MiD than the other two surveys. This could be because MiD makes it easier to record business trips all at once without needing to provide individual entries in the travel diary.

From figure 8 it can be observed that trips with an unknown purpose include virtually no active travel. This supports the case for disregarding trips with unknown purposes, which is consistent with transport modelling convention.

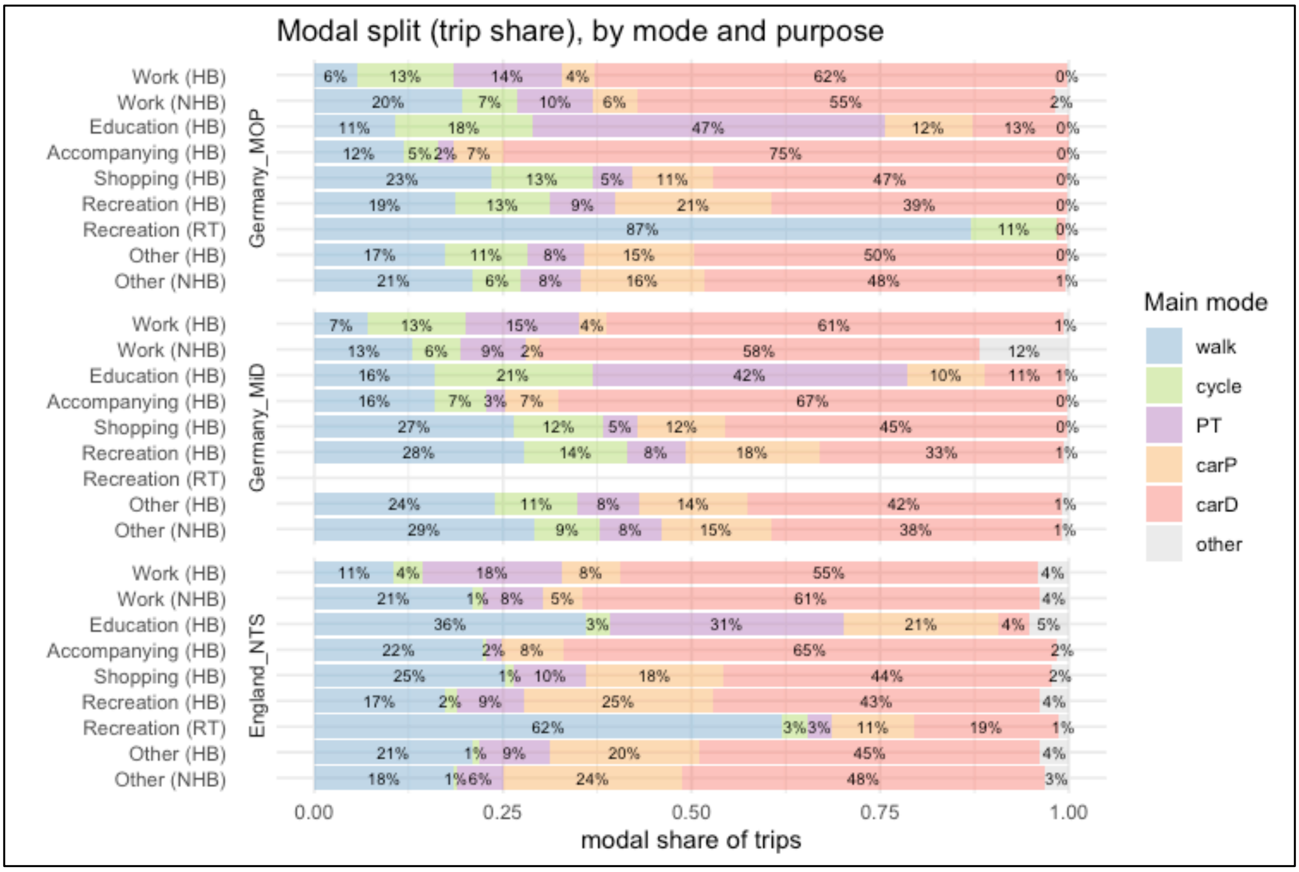


Figure 7: Modal split (trip share), by mode and purpose

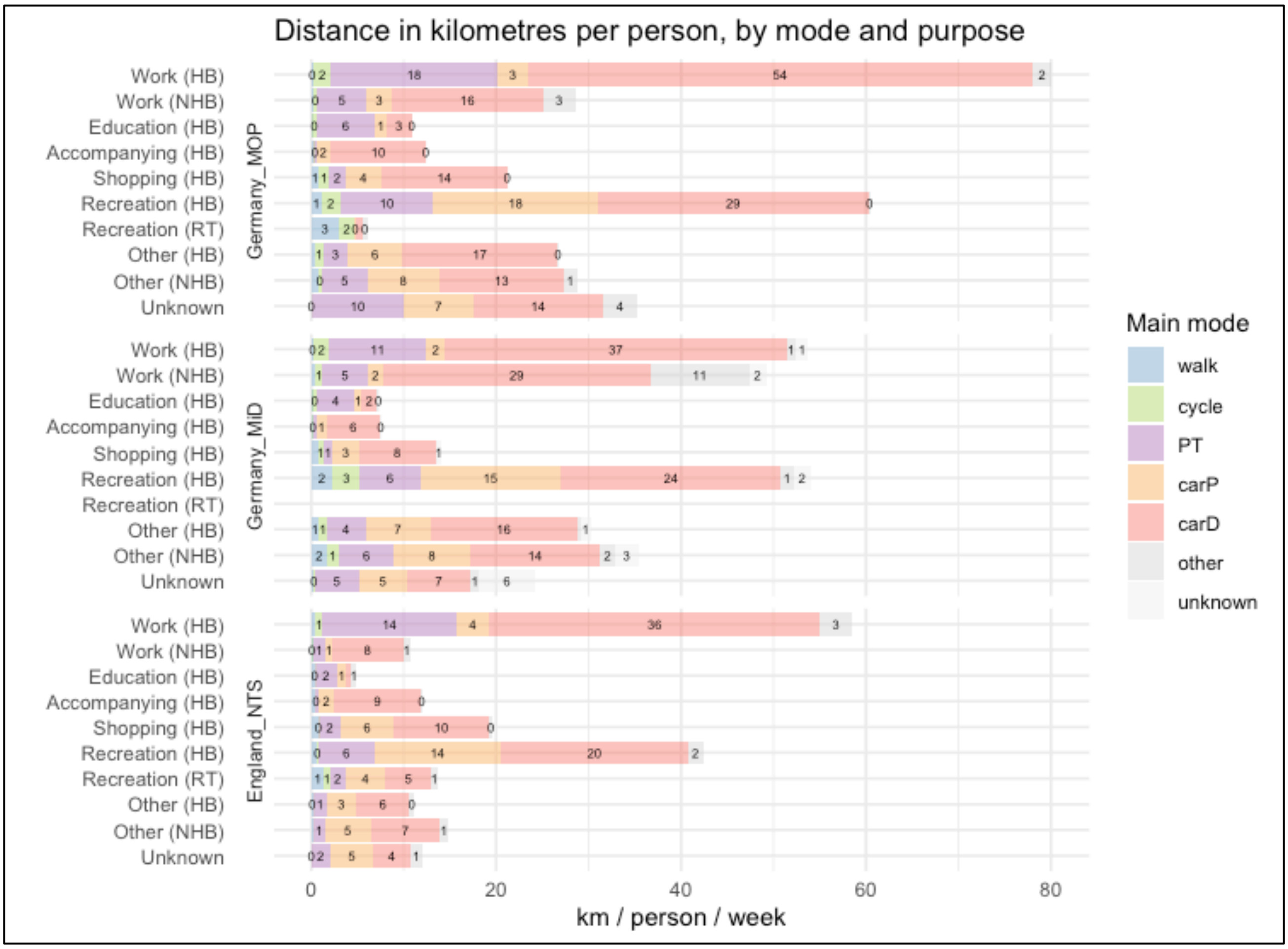


Figure 8: Distance in kilometres per person, by mode and purpose

4.2. Physical Activity Assessment

Figure 9 shows the assessed transport-domain PA by mode. Recall from section 3.3 that only PA from walking and cycling is considered. However, if walking and/or cycling are part of a multi-modal trip, then this PA is also captured even if the main mode is different.

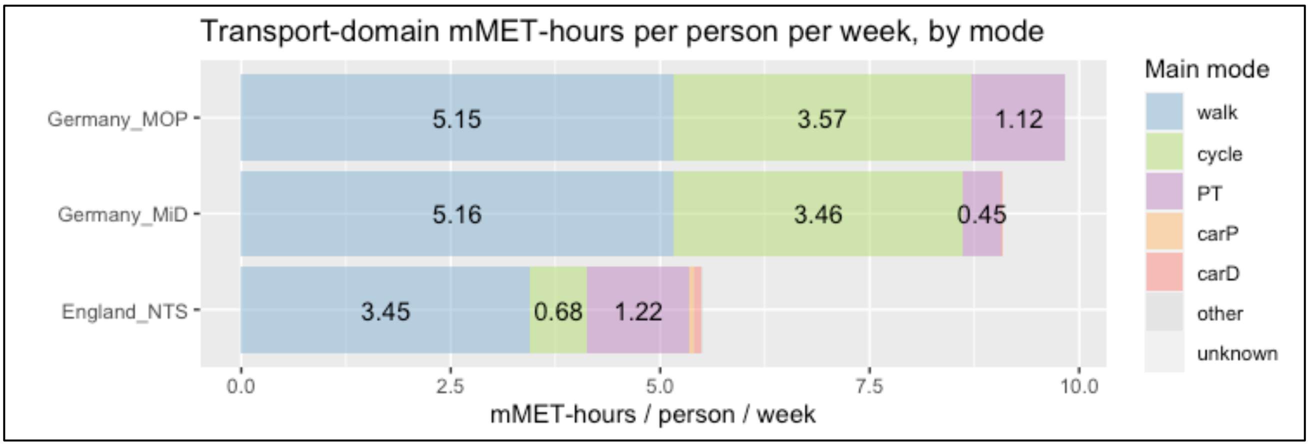


Figure 9: Transport-domain mMET-hours per person per week, by mode

Figure 9 claims that Germans are about twice as active as the English in the transport domain. On average, walking and cycling trips in England account for about 46% of the recommended 8.75 mMET-hours per week, whereas in Germany it is about 95%. These results seem inconsistent with comparisons for all-domain PA. For example, a 2018 WHO report [77] that considered all-domain PA found that the proportion of adults who achieve guideline PA levels is actually higher in England (67%) than in Germany (46%). Comparing the two German datasets, PA captured from walking and cycling trips was almost exactly the same, which is a useful validation of the MOP.

Regarding PT access and egress, none of the datasets captured as much PA (relative to other modes) as the objective study by Chaix et. al. [46]. This suggests that all of the assessed HTSs underestimate PA from PT. The British NTS captured the most PT PA, despite having the fewest PT trips overall (recall figure 5). This suggests that using a stage-level approach is the most effective way to capture PA in PT trips. Comparing the two German datasets, PT PA was significantly higher in the MOP than NTS. This is probably because the MOP included information about work- and education-end PT distances but the MiD did not.

Differentiation by Age and Sex

Figures 10 and 11 compare transport-domain PA by sex and age. Figure 10 shows that males are generally more active in transport, although the difference between males and females is almost entirely attributed to higher cycling uptake among males (1.5x in Germany and 4x in England). Walking is slightly lower among males, which is probably because of substitution between walking and cycling.

Figure 11 shows that the relationship between active travel and age is different for the UK than for Germany. In the UK active travel is highest among the younger age groups and decreases gradually with age. However, in Germany, both the MiD and MOP reported that active travel is highest among those aged 60–69, followed by those aged 50–59. People in their 20s had the lowest transport-domain PA. This contradicts a recent GPAQ study by Wallmann-Sperlich and Froboese [78] which found that transport-domain PA in Germany was highest for this age group. Nevertheless, the differing patterns in England and Germany reflect patterns in overall physical activity as reported by the WHO: in Germany, the proportion of sufficiently active adults declines only slightly for those age 65 and above (from 46% to 42%), whereas in England the decline is more substantial (From 67% to 44%).

The differences between age groups are more dramatic in the German MOP than in MiD, but the overall pattern was the same. This could be because of the much smaller survey population in the MOP data, causing a sampling bias.

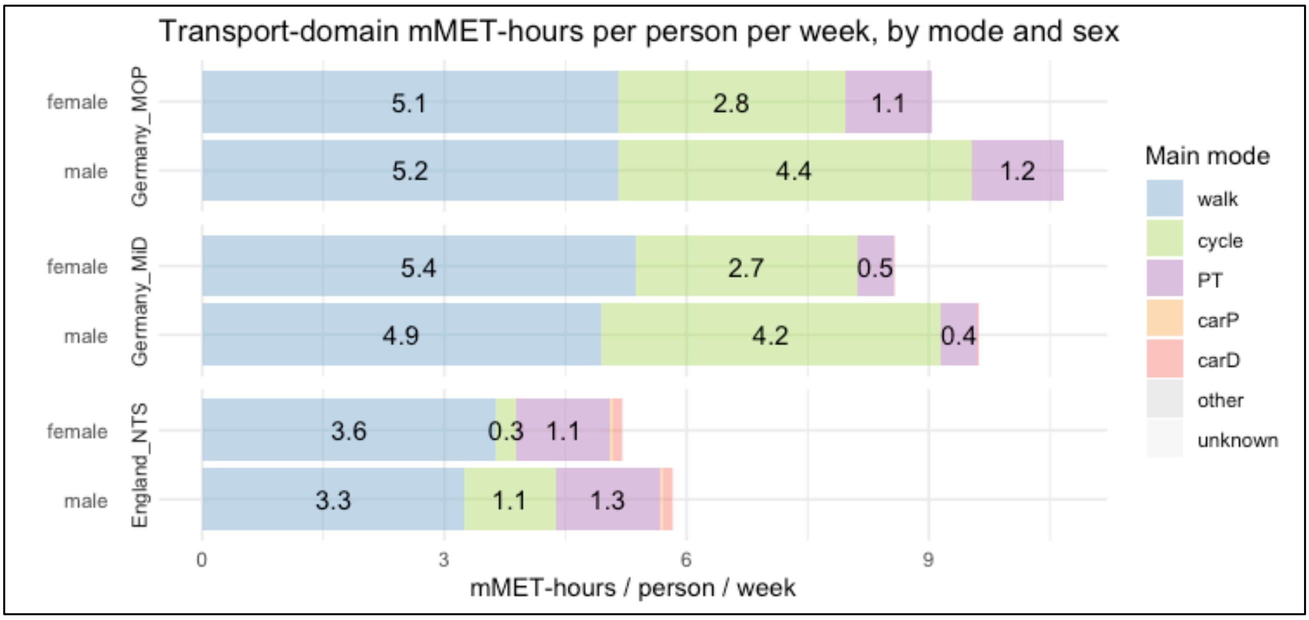


Figure 10: Transport-domain mMET-hours per person per week, by mode and sex

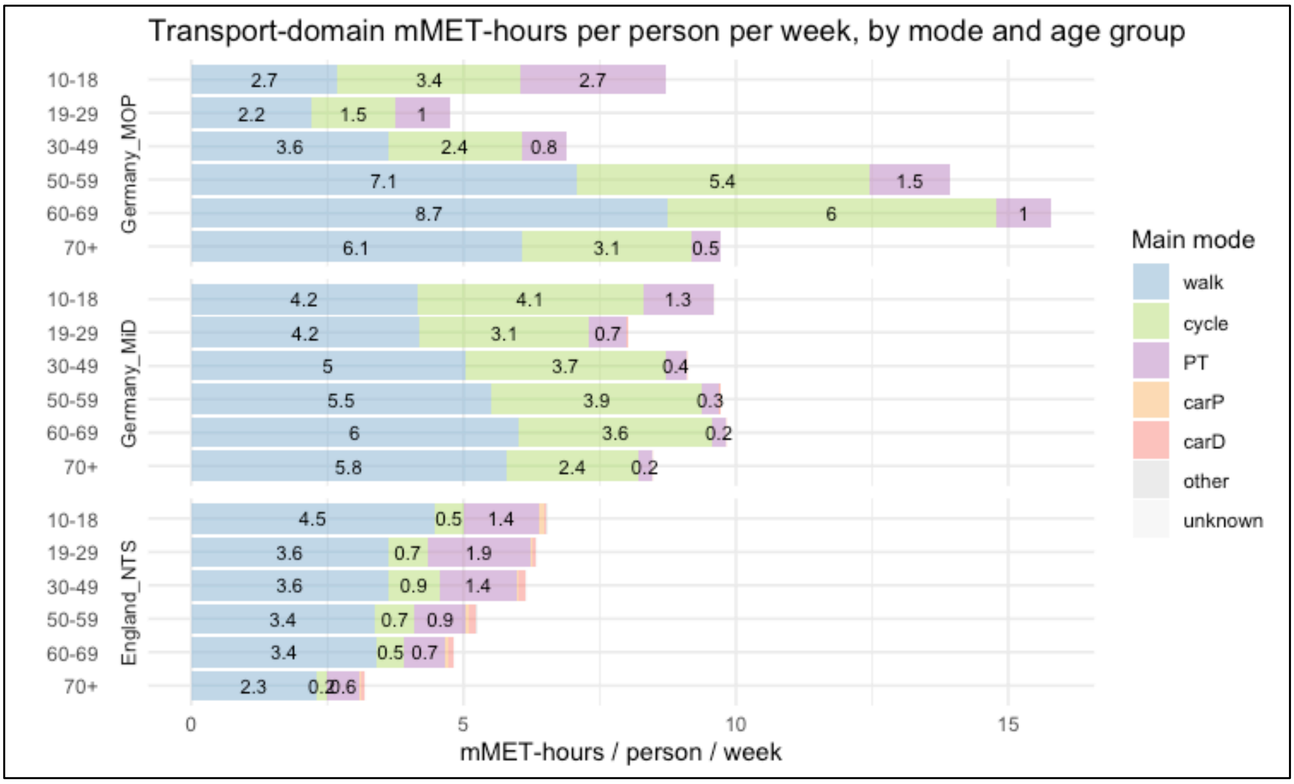


Figure 11: Transport-domain mMET-hours per person per week, by mode and age group

Differentiation by Purpose

Figure 12 shows how transport-domain PA is broken down by purpose. It reveals the difficulties in harmonising trip purposes across datasets. Recall from table 4 that trip purposes were similar for both the MOP and NTS, but the MiD purposes were substantially different: MiD was also the only survey that did not differentiate between recreational roundtrips and other recreation trips; it was also the only survey that differentiated between business trips and other work trips. The following key observations are made:

- For the MOP and NTS, recreational round trip was by far the most active purpose. In the MiD these trips would have instead coded as either ‘Recreation (HB)’ or ‘Unknown’ if they started at home, ‘Work (NHB)’ if they started at work, or ‘Other (NHB)’ if they started elsewhere. This explains why those purposes have more PA in the MiD than in the other surveys.
- Active travel for purpose ‘unknown’ is negligible in MOP and NTS but not MiD. This is possibly due to some MiD recreational roundtrips falling into the ‘unknown’ purpose, as described above.
- NHB purposes are very low in the MOP and NTS, but high for MiD. Again, this could be due to the presence of trips that would have otherwise coded as recreation roundtrips. For the ‘Work (NHB)’ purpose, this could also partially be attributed to the way business trips are recorded, described in section 4.1.

Because of the significance of recreational roundtrips in transport-domain PA, it would appear important not to neglect them in transport models if the goal is to model health impacts. In addition, because of the unique modal split properties of recreational roundtrips (see figure 7), the purpose would ideally be defined separately from other purposes. This supports the use of survey data that, unlike MiD, has this purpose defined explicitly in the data.

PA in PT trips was well captured for HB work trips in both the MOP and NTS. For English work commutes, more PA occurs though PT than through main-mode walking or cycling. In the German MOP, PT PA was highest for HB work and HB education. Not coincidentally, these were the only MOP purposes in which it was possible to obtain PT access and egress distances on both ends of the trip.

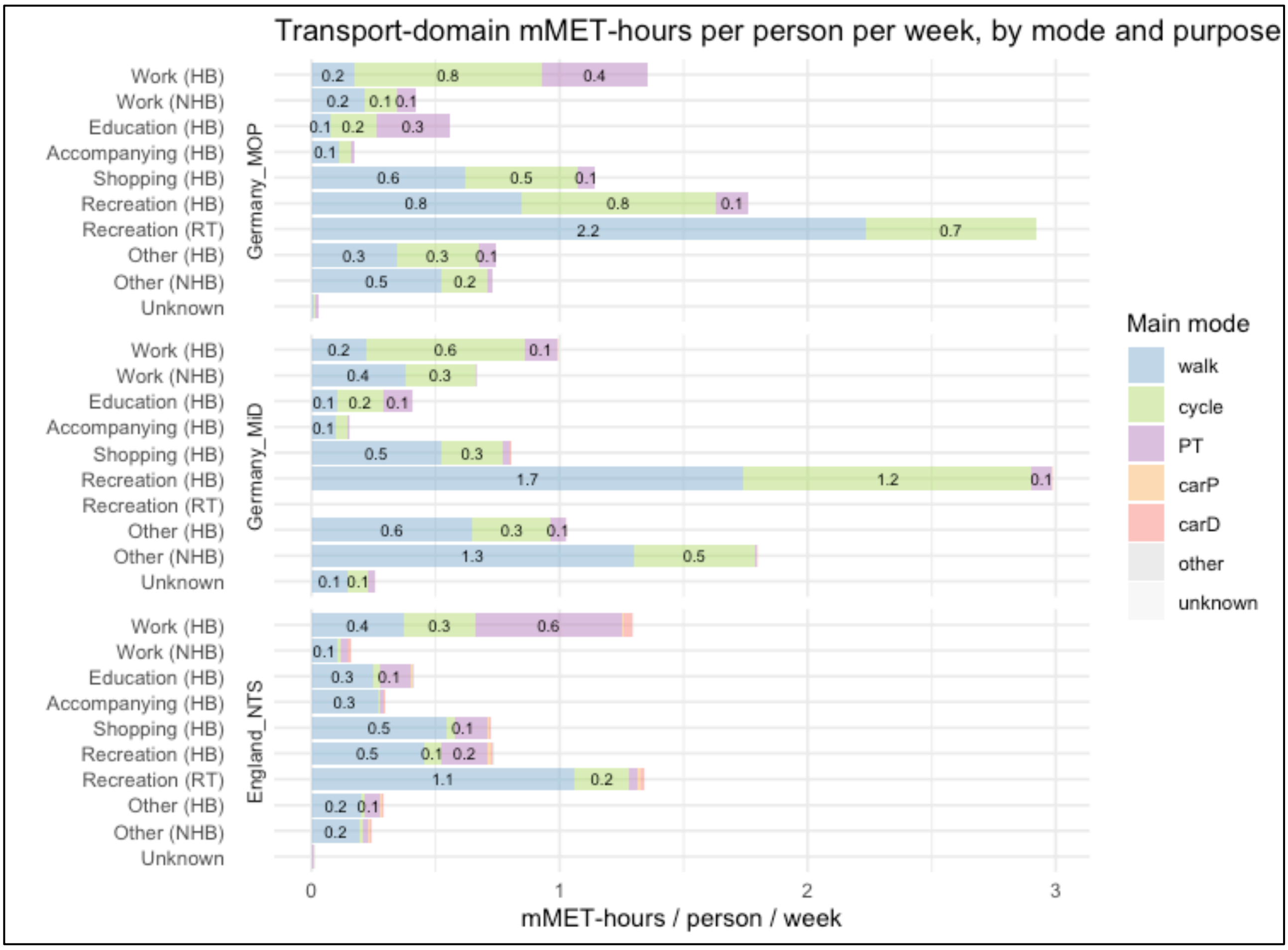


Figure 12: Transport-domain mMET-hours per person per week, by mode and purpose

4.3. Discussion

This cross-dataset comparison reveals how some HTS methods are better suited for PA assessment than others:

The British NTS appears most effective at capturing PA from PT access and egress, thanks to its stage-level reporting. However, the NTS has an inexplicably low trip reporting rate, which is possibly causing a systematic underestimation of transport-domain PA.

The German MiD has strengths in its size, scope, and precision. However, it is limited from a PA perspective because it poorly captures PT access and egress and does not differentiate recreational roundtrips, which are a major component of PA.

The German MOP's key strength is its full 7-day travel diary. However, it is still relatively poor at capturing PT access and egress, and does not clearly differentiate mobile from non-mobile individuals. Furthermore, because of its small sample size, age-specific findings in the MOP appear to be unreliable.

Chapter 5. Agent-level Assessment using the *Deutsche Mobilitätspanel*

Of the three datasets included in this study, only with the MOP was it possible to make observations about the weekly transport-domain PA distribution at the individual level. Using this dataset, transport-domain PA could be calculated for each individual by combining the mMET-hours captured through in each trip over the 7-day period. This made it possible to investigate the distribution of transport-domain weekly PA volume across the population.

Table 5 provides a summary of this distribution. It shows the proportion of individuals who meet specific mMET-hour thresholds, depending on the types of PA that were considered in the calculation. The 0 mMET-hour threshold shows the proportion of individuals who perform any transport-domain PA. The 8.75 mMET-hour threshold shows the proportion of individuals who meet WHO minimum guideline PA levels through transport-domain PA. The 17.5 mMET-hour threshold shows the proportion of individuals who meet the WHO guideline for 'additional health benefits' [8, 45] through transport-domain PA. This table also shows mean and max mMET-hours per week across individuals.

Table 5: Distribution of individual weekly transport-domain physical activity (summary)

Included Physical Activity Types	Percentage over 0 mMET-hrs	Percentage over 8.75 mMET-hrs *(WHO minimum)*	Percentage over 17.5 mMET-hrs *(WHO 'additional health benefits')*	Mean (mMETs-hrs)	Max (mMETs-hrs)
Walk (main mode) Cycle (main mode) PT (home-end) PT (work/edu.-end)	81%	41%	19%	10.11	140.04
Walk (main mode) Cycle (main mode) PT (home-end)	81%	38%	17%	9.49	140.04
Walk (main mode) Cycle (main mode)	77%	34%	16%	8.75	140.04

Note that the mean values in table 5 in are similar to the MOP population means calculated previously in figure 9. However, they do not match exactly because of person-weights applied in this individual analysis that could not be applied in the aggregate analysis.

Table 5 reveals that most Germans aged 10+ (77%) use an active mode as their main mode at least once during the week. If PT access and egress is included, this rises to 81%. The table also shows that 41% of Germans aged 10+ meet guideline recommendations through PA in the transport domain alone. This is very close the WHO's value of 46% of adults, which was calculated based on all-domain PA [77].

The full distribution can be visualised using the empirical cumulative density function (ECDF) as shown in figure 13. In this plot, the y-axis corresponds to the percentage of individuals who have achieved weekly MET-hours over the value in the x-axis. Therefore, the values at $x = 0$, $x = 8.75$ and $x = 17.5$ (dashed) correspond the summary values in table 5.

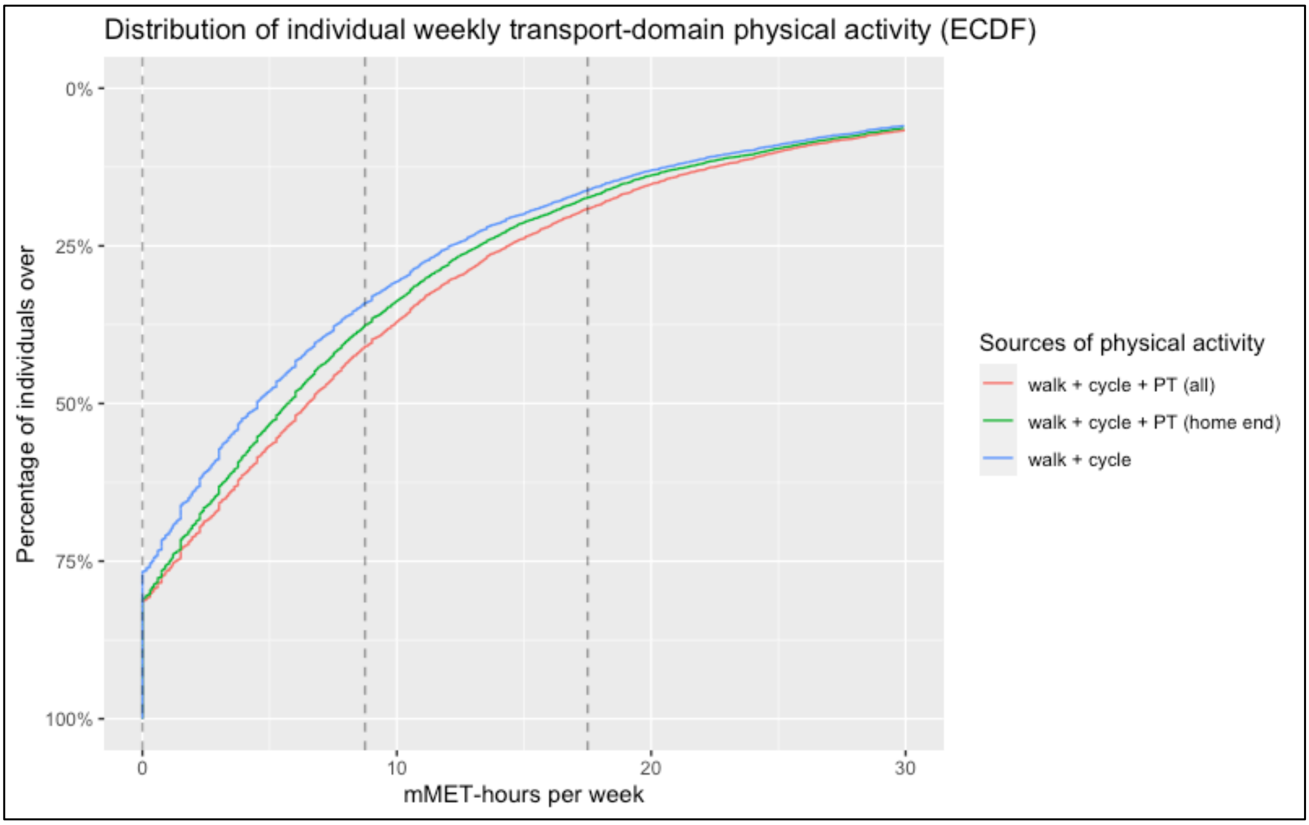


Figure 13: Distribution of individual weekly transport-domain physical activity (ECDF)

ECDF plots like the one shown in figure 9 are especially useful for comparing the population distribution of PA under different circumstances. As this is the primary indicator for this thesis, several of these plots will be shown throughout.

From the literature review it became clear that PA access and egress is an important component of transport-domain PA and should be considered. Using the MOP, it was possible to capture PT access and egress at the home, education, and work trips ends. However, there was no information about shopping, recreation and 'other' trip ends. Assessing PA using this 'incomplete' access and egress data would be biased towards individuals who take more work/education trips, as opposed to individuals who are part-time or unemployed and might take more trips of other purposes. To prevent such bias, the remainder of this study includes only PT access on the home end of trips (i.e. the green line in figure 13). It is acknowledged that this underestimates the health benefits of PT.

5.1. Defining Physical Activity Levels

To understand the correlates and determinants of physical activity, the MOP population was broken down into 5 pseudo-quintiles based on their level of transport-domain PA. These are shown in figure 14. The lowest level ("inactive") was assigned to the 19% of individuals with no transport-domain PA. Levels "low" and "medium" are evenly-split groups of individuals who are active but do not meet the WHO minimum guideline limit of 8.75 mMET-hours per week. Individuals at level "high" are between this guideline limit and the 'additional health benefits' guideline of 17.5. Individuals at the level "very high" exceed the 'additional health benefits' guideline.

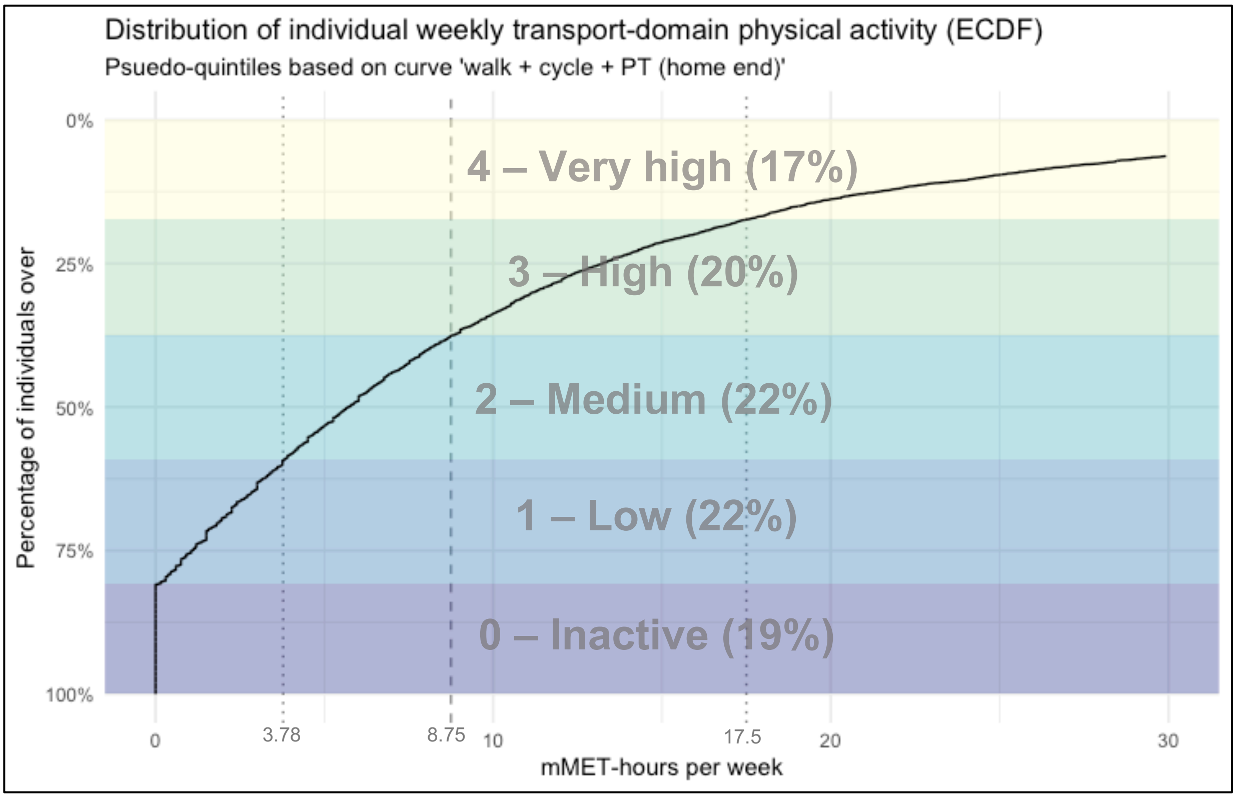


Figure 14: Psuedo-quintiles based on ECDF for transport-domain physical activity

Figure 15 shows the average total distance travelled on each mode by individuals in each PA category. Figure 16 shows the average total number of trips travelled. These figures reveal that individuals who are more active generally take more trips but travel less distance.

As expected, active modes shares increase as PA levels increase. However, this is dominated by increases in cycling, especially at the more active levels. Increases in walking are relatively steady in comparison.

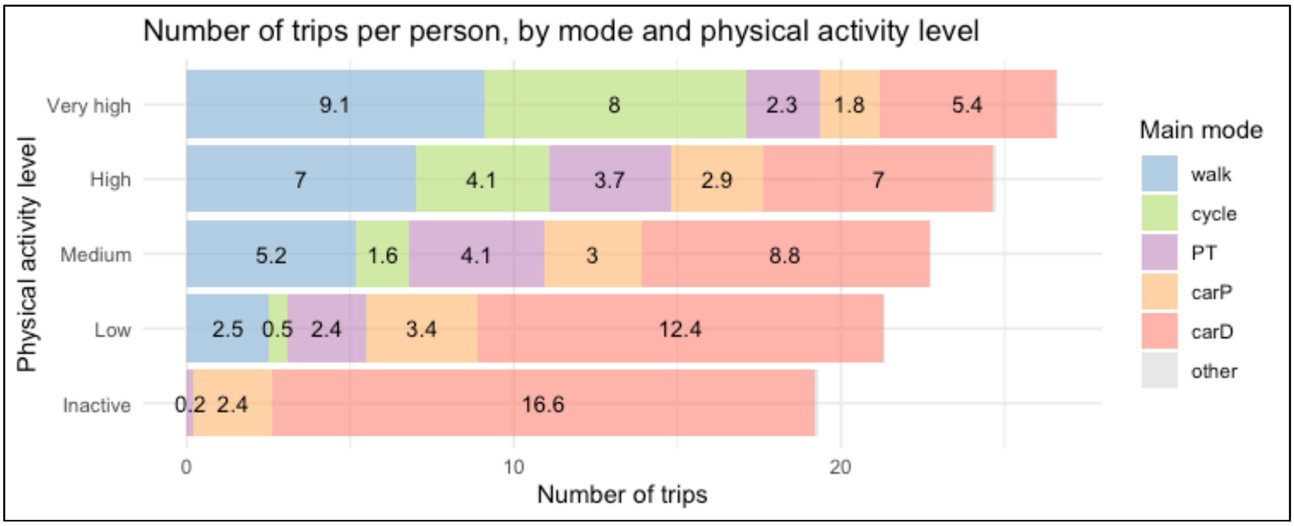


Figure 15: Number of trips per person, by mode and physical activity level

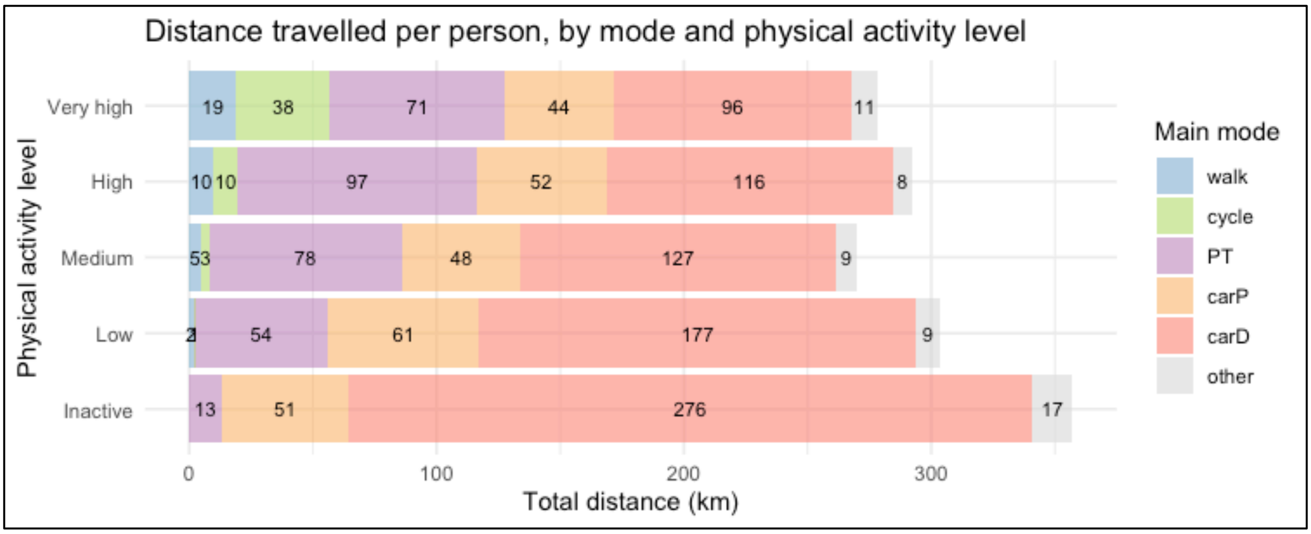


Figure 16: Distance travelled per person, by mode and activity level

Another important observation can be made regarding mode substitution. As individuals become more active, active modes replace car driver more than they replace any other mode. The mean car driver kilometres travelled drops by 65% between the inactive and most active groups (from 276km to 96km). This shows that individuals' transport carbon footprint drops substantially as people become more active. Car passenger and PT have a different pattern: the share of these modes is highest for people in the middle categories. Their share is second lowest for completely inactive people (where travel is dominated by car driving), and lowest for the most active people (where travel is dominated by active transport).

Figure 17 shows the proportion of individuals in each category who use each mode at least once during the week. In other words, it represents the proportion of individuals who can be identified as walkers, cyclists, PT users, and so on. It shows that the proportion of walkers only increases slightly between the lowest and highest level (76%–88%), but the proportion of cyclists increases substantially and steadily, from 16% in 'low' to 70% in 'very high'. The proportion of PT users is highest for individuals at that 'high' level, but drops off at 'very high', probably because of substitution toward cycling. Of those individuals who exceed WHO guidelines (i.e. ‘high’ and ‘very high’), the majority are cyclists. From a policy perspective, this pattern is intuitive: it means that increasing the proportion of people who are walkers would primarily reduce the number of people who are inactive (in the transport domain). On the other hand, increasing the proportion of people who are cyclists would more equally benefit people at any level. The likelihood of being a car driver decreases as physical activity level increases, but it remains a majority.

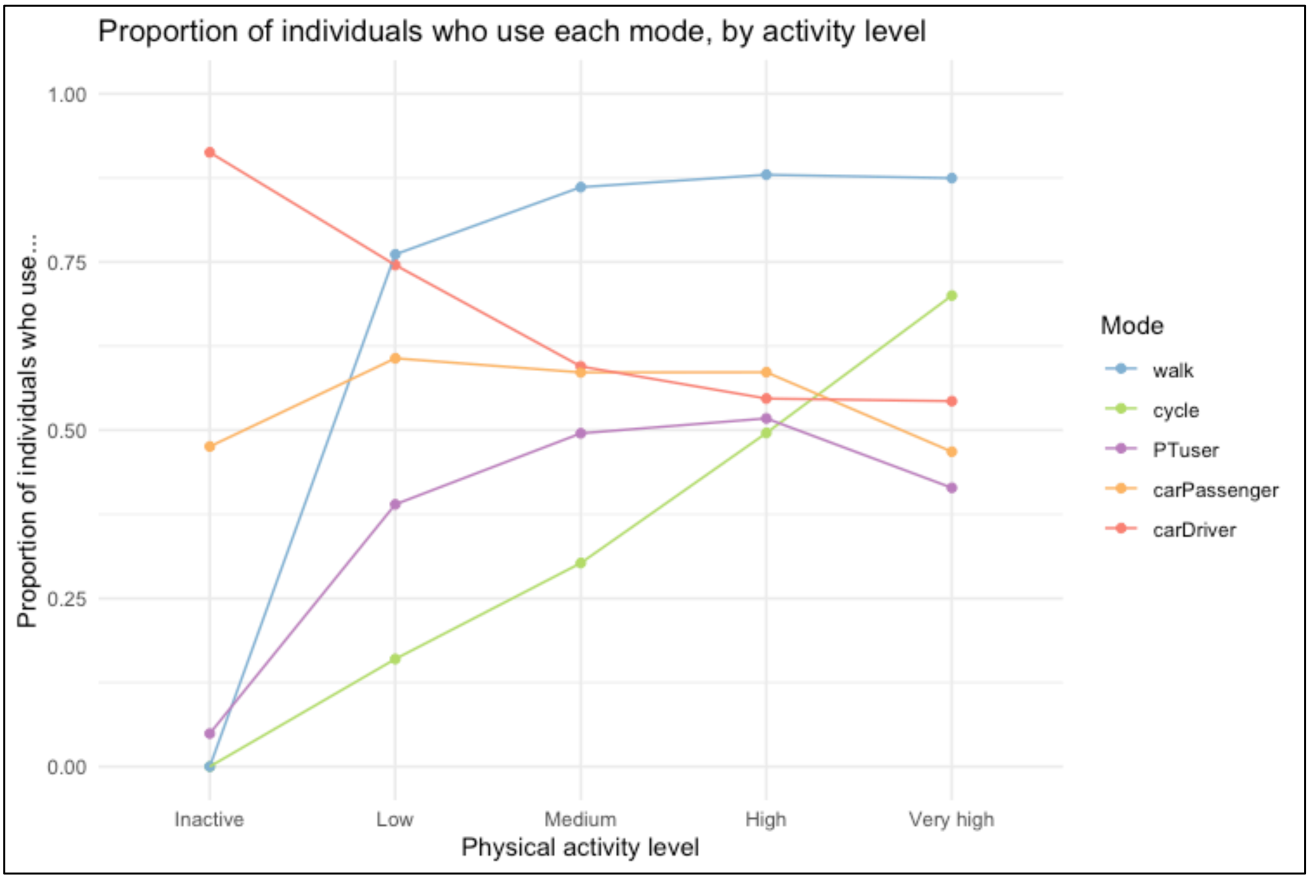


Figure 17: Proportion of individuals who use each mode, by activity level

Figure 18 shows the number of trips per person in each PA level, differentiated by mode and purpose. It is similar to figure 15 but with a further differentiation by purpose. The proportion of each mode along a bar is equivalent to the modal split for this category. This figure shows that for recreational roundtrips, the change in activity level has a significant impact on the number of trips. However, for mandatory purposes (work & education), it the difference in activity level impacts the share of active modes rather than the total number of trips. For shopping and recreation, both mode share and number of trips are affected.

This figure suggests that having more work trips makes individuals more likely to be inactive. A separate analysis of work trip distances (not shown) also found an inverse relationship between work distance and activity level. The relationship between the work trips and PA will be explored further in sections 5.2 and 6.2.

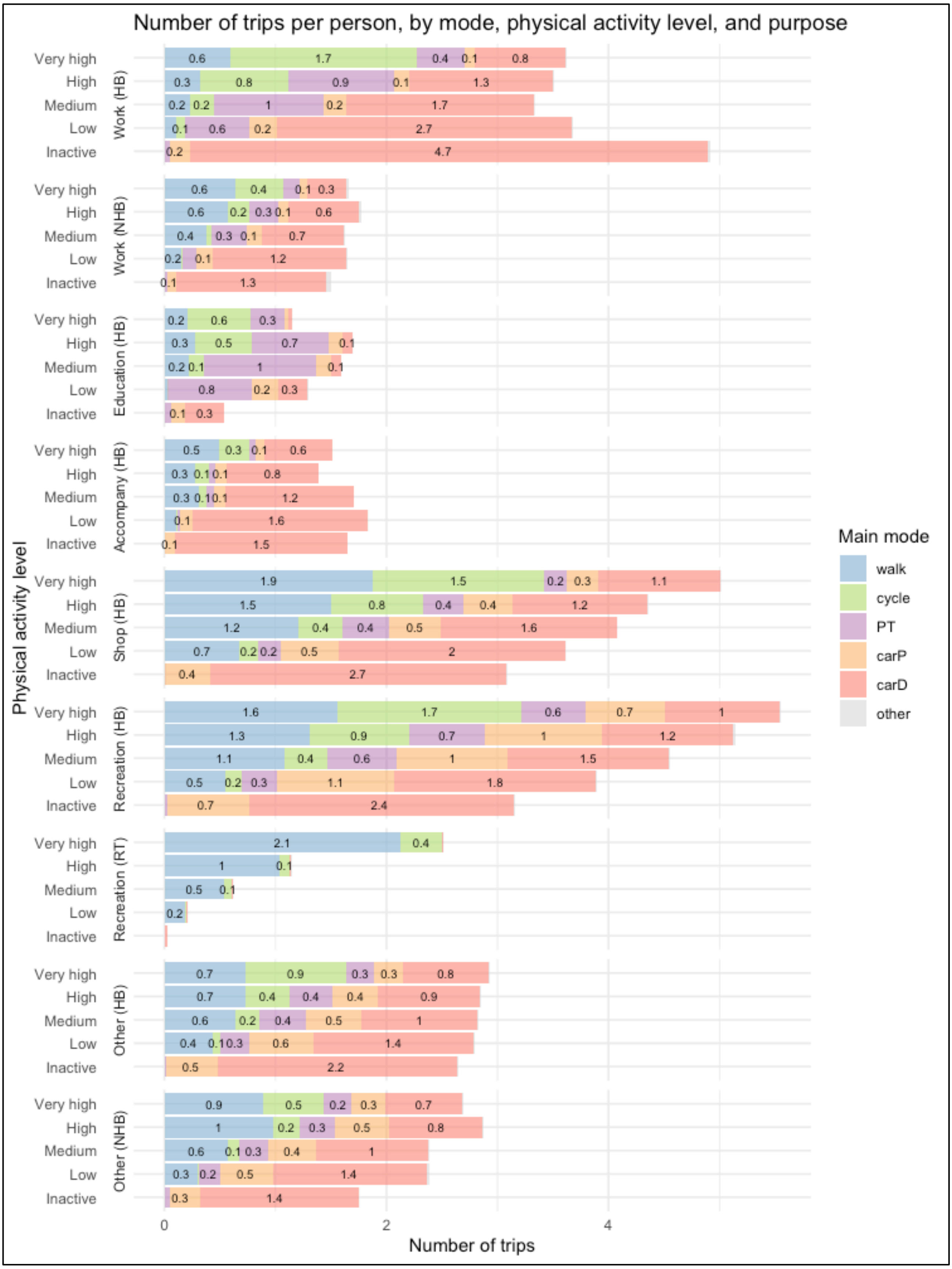


Figure 18: Number of trips per person, by mode, physical activity level, and purpose

Figure 19 shows the relationship between activity level and sex and age. It shows that females are more likely to be active (i.e. not inactive) in transport than males, despite their total PA being lower, as shown previously in figure 10. Only the highest PA category has more males than females, which is likely a result of increased cycling uptake among males. Regarding age, children tend towards the middle PA levels (low, medium, and high). Individuals aged 19-39 tend to be less physically active. Those aged 40–69 tend towards either extreme (either completely inactive or very active). Finally, individuals aged 70+ tend towards the central levels. The differing characteristics of each age group align with the aggregate results given previously in figure 11, which showed that average total PA was lowest for younger adults, highest for those aged 60–69, and in the middle for children aged 10–18 and adults aged 70 and up.

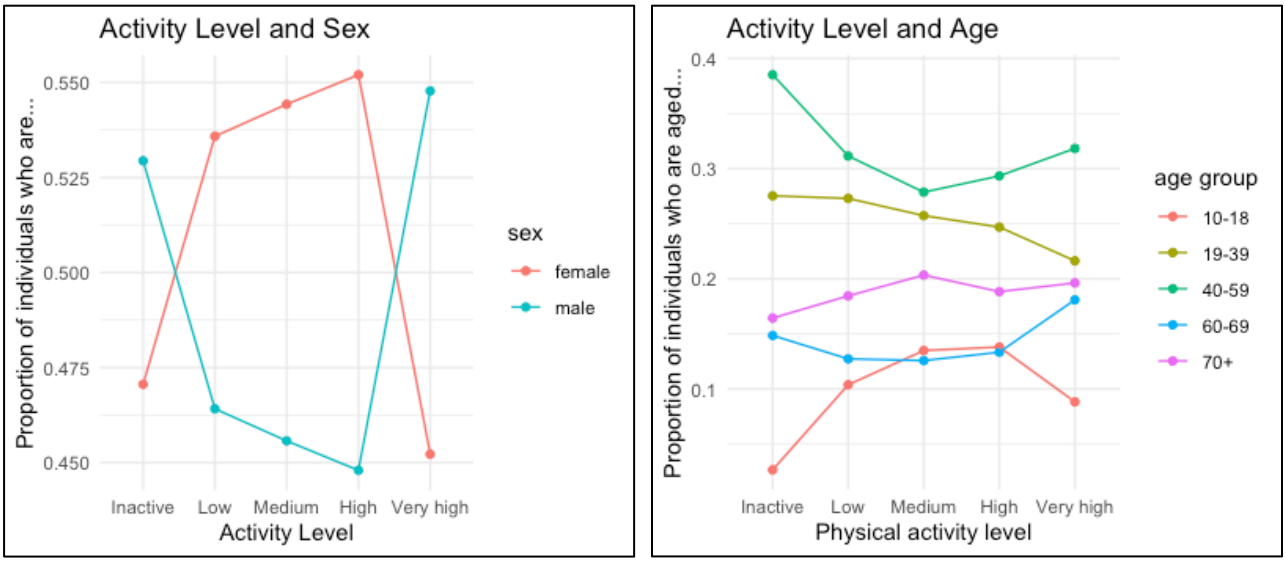


Figure 19: Activity level and sex (left) and age (right)

Figure 20 shows how mMET-hours of individuals in each category are broken down into short walks, other walks, cycle journeys, and PT access trips. Short walks are defined here as walks less than or equal to 1.6km, similar to the definition used in the British NTS. These breakdowns were calculated by first calculating the mMET-hour split for each individual, and then calculating an average split across individuals in each activity group.

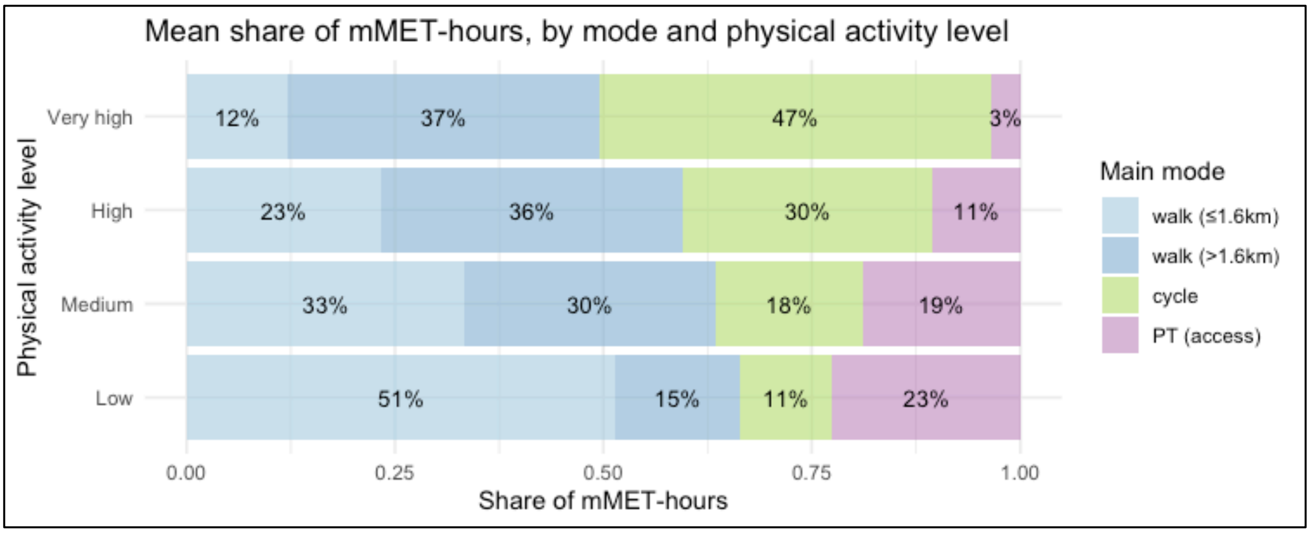


Figure 20: Mean share of mMET-hours, by mode and physical activity level

Short walks dominate for individuals in the 'low' PA level, making up on average 51% of their mMET-hours. This and PT access combined make up 74% of the mMET-hours of low-activity individuals. This reveals the importance of short walks and PT to people who are only slightly active in their travel. In fact, 65% of individuals in the low-PA category take *only* short walks or PT access trips (not shown). If short walks were to be neglected, or only captured once per week, most of the PA for these 'low' activity individuals would be missed, leading to an underestimation of the number of individuals in that category. Similarly, if only one seventh of these short walks were to be captured once and multiplied by a factor of 7 (as in the British NTS), then the individuals who took these walks would have an artificially high PA level. For this reason, the British NTS would not be suitable for this type of agent-level investigation.

5.2. Impact of Mandatory Travel Patterns

This section uses investigates how mandatory travel influences discretionary travel and therefore PA. This analysis considers two properties of individuals' mandatory travel behaviour, based on the adapted modelling framework described in section 3.6. The first property is the type and number of commute trips. The second property is the dominant commuting mode, defined as the mode used most frequently over the week for these trips. Individuals who commute for both work and education were excluded from this analysis because of the insufficient sample size (only 2% of individuals).

Commute type and frequency

Individuals were differentiated according to their commute type (worker, student, or neither) and commute frequency (number of trips). Commute frequency is defined as the number of round trips from an individual's home to their place of work or study. Some frequencies were combined to achieve a reasonable sample size (minimum 300) for each group.

Before continuing with this analysis, it is important to note that there is a strong relationship between commute type, frequency, and age, as shown in figure 21. Individuals with no mandatory trips are mostly older and retired, while individuals with education commutes are mostly children. This makes it difficult to draw comparisons across commuter types (i.e. students vs. workers vs. neither) independently of age. The issue would ideally be addressed by re-weighting individuals to control for age; however, due to the small sample size this is not plausible. This is taken as a limitation of this study.

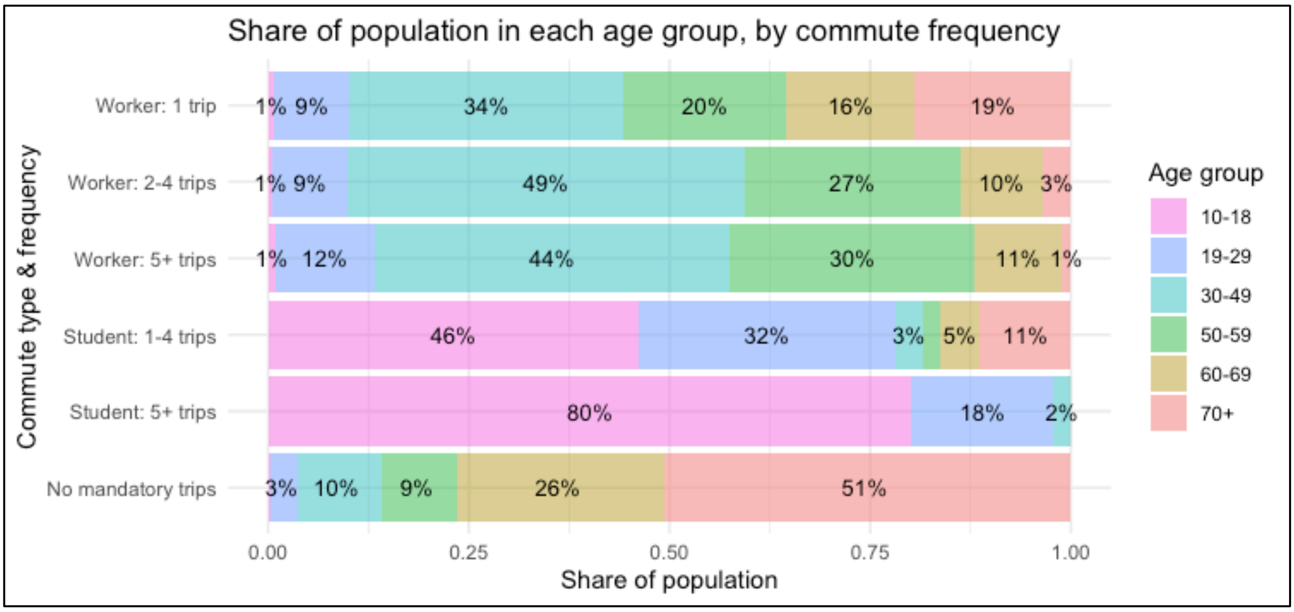


Figure 21: Share of population in each age group, by commute frequency

The relationship between commute frequency, trips, and modal split is shown in detail in figure 22. For each commute type, this figure plots the number of trips taken for each purpose and mode. This plot only counts the outbound portion of HB trips, which is an approach consistent with trip-based modelling. The modal split (trip share) for each category is written. The plot shows that for workers, more work trips correspond to less discretionary trips, with the exception of NHB work trips. For students, the pattern is similar; however, students have more HB recreation, HB other, and NHB other trips than workers in general. Individuals who are not workers or students take the most shopping and recreational round trips. They also take more HB recreation trips than workers, but not as many as students. Regarding mode share, figure 22 shows that on average, workers with more commute trips are more likely to choose active modes for their commute, but less likely to choose active modes for their discretionary trips. This could possibly be explained by the fact that these individuals are spending more time travelling for work, and therefore placing a high importance on travel time for their discretionary trips in order to stay within their travel time budget. For students, more education trips correspond to a higher share of active modes for their education trips. However, there is no clear relationship between commute type and active mode share for students' discretionary trips.

Figure 23 shows how this relationship translates into PA. As expected, individuals with more mandatory trips are more active for mandatory trip purposes. However, they are less active in other purposes. If the PA volumes across all purposes are added up (not shown), then on average, the gain in PA for taking more mandatory trips more than compensates for the loss in PA for taking less discretionary trips. However, as figure 24 shows, this gain is not distributed evenly in the population. Having more work commutes causes workers to tend towards either extreme. Workers with 5+ work trips are more likely to be among the most active individuals, but they are also the most likely to be inactive. This pattern is not the case for students, who are generally less inactive then workers. This may be because students are generally children and, as figure 19 showed earlier, children tend towards the 'medium' levels of PA. Besides age, there may also be many other confounding factors at play, such as income and the built environment. These will be considered during model development in chapters 6 and 7.

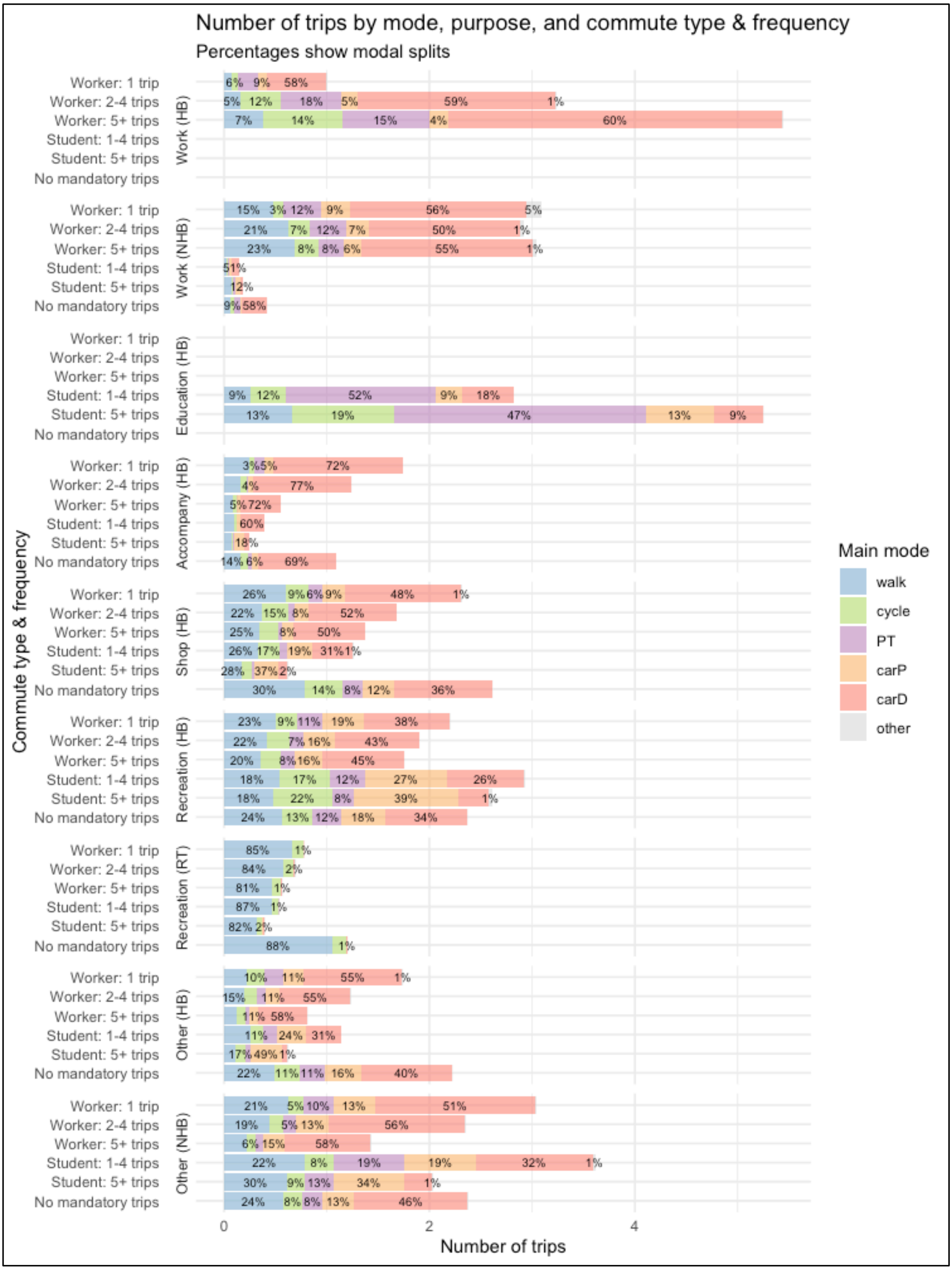


Figure 22: Number of trips by mode, differentiated by purpose, commute type, and frequency

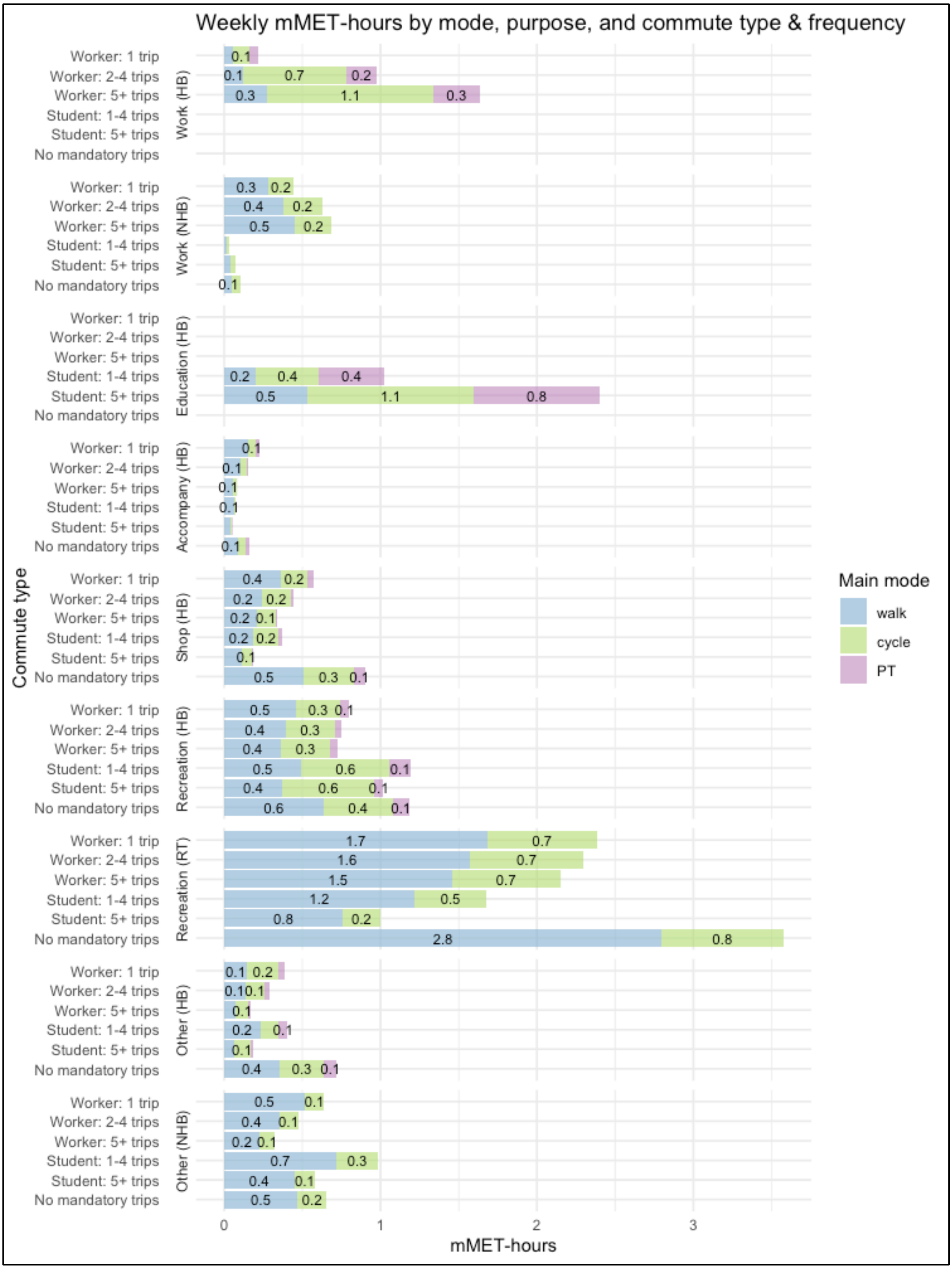


Figure 23: mMET-hours by mode, differentiated by purpose, commute type, and frequency

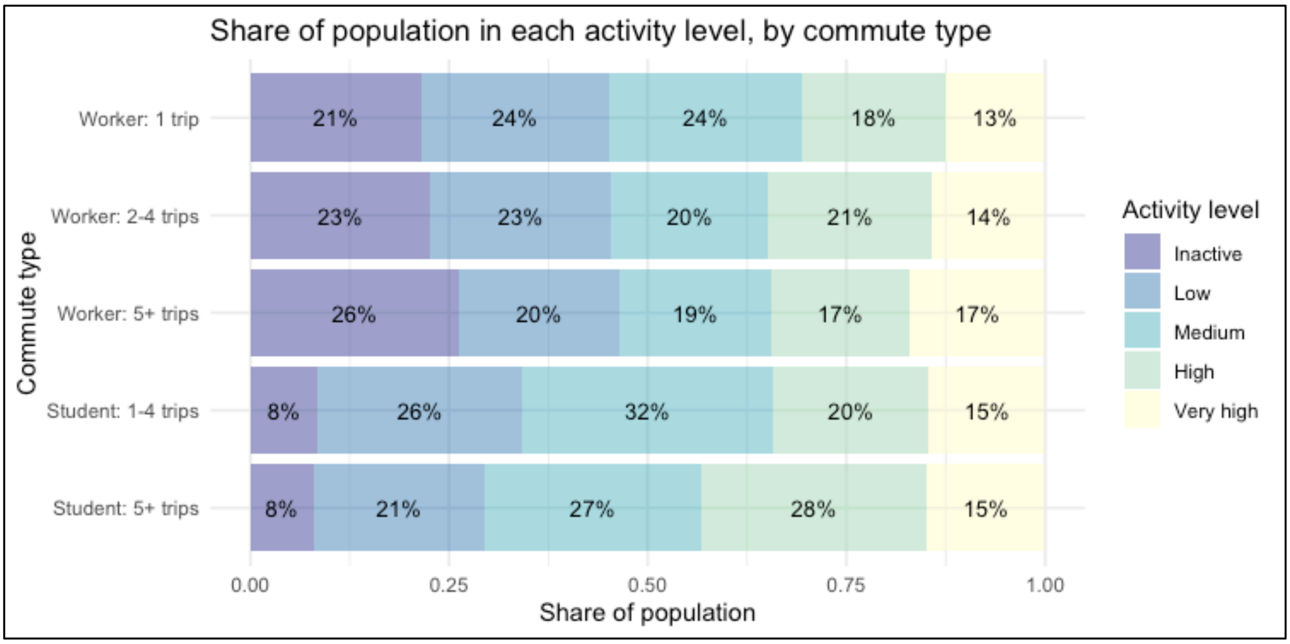


Figure 24: Share of population in each activity group, differentiated by commute frequency

Primary commute mode

This analysis is similar to the previous analysis on number of mandatory trips, except that individuals will now be differentiated based on their primary mandatory mode (i.e. their usual commute mode) instead of commute frequency. As introduced in section 3.6, this is defined as the mode used most frequently for an individual's mandatory trips throughout the survey week. If there is a tie, it is the mode that accounts for the most travel time.

First, figure 25 relates this analysis to the previous analysis on commute frequency. The share of individuals based on their usual commute mode is similar to the mode share of trips shown previously in figure 22. It suggests that workers with more frequent work trips are more likely to choose active modes over car passenger and PT as their dominant mode. However, the proportion of car-driving commuters remains the same regardless of commute frequency. Students strongly favour PT over private car, likely because of their age. They are also more likely to have active dominant modes. Part-time students are more likely to drive than full-time students, likely because they are generally older and have a higher income.

Figure 26 shows the relationship between usual commute mode and number of trips for each mode and purpose. Individuals who usually commute by PT or as a car passenger generally travel to work fewer times per week, which is consistent with the findings in figure 25. In fact, car passenger and PT commuters generally take the fewest mandatory *and* discretionary trips. Regarding discretionary modal splits, non-commuting individuals (labelled with "N/A") tend to favour walking. Commuting individuals consistently favour using their dominant commute mode for their discretionary trips. Car-driving commuters are generally the least likely to choose active modes for their discretionary trips. The exception to this is recreational roundtrips: car-driving commuters generally take more recreational roundtrips than car-passenger or PT commuters, perhaps as a form of compensation to introduce PA into their car-oriented lifestyle. Nevertheless, active commuters still take the most recreation roundtrips.

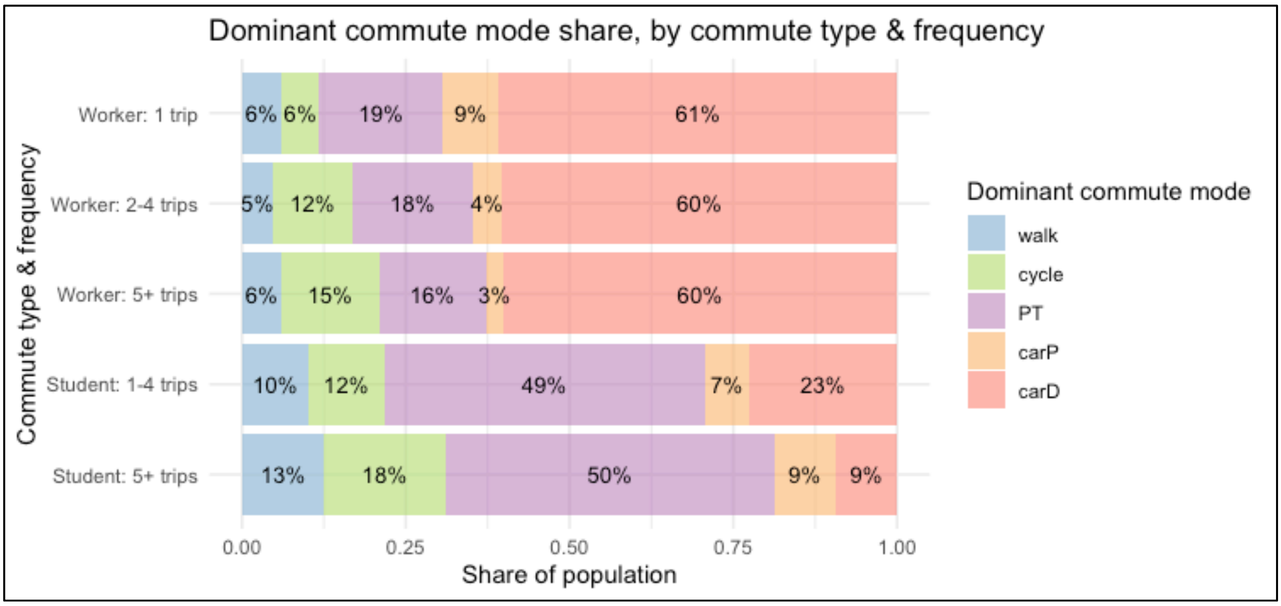


Figure 25: Population shares based on usual commute mode, by commute type

Figure 27 translates this into PA volume. The pattern is similar to the pattern with total trips, with car-driving commuters gaining the fewest mMET-hours for all purposes except recreational round trips. Cycle commuters, despite taking a similar number of work trips as walking commuters, gain around double the mMET-hours in their work commutes as result of commuting much longer distances. Cycle commuters also gain the most mMET-hours in other purposes, with the exception of accompanying (highest for walking commuters) and recreation roundtrips (highest for non-commuters). If mMET-hours are added up across all purposes (not shown in the figure), we find that car-driving commuters are generally the least active (mean 4.6 mMETs) while cycling commuters are the most (mean 21.8 mMETs). Given this information, the distribution of mMET-hours based on commute mode, shown in figure 28, is not surprising. Most cycle-commuters have 'very high' transport PA, and most walkers are 'high' or 'very high'. On the other end, car drivers are most likely to be completely inactive, followed by car passengers.

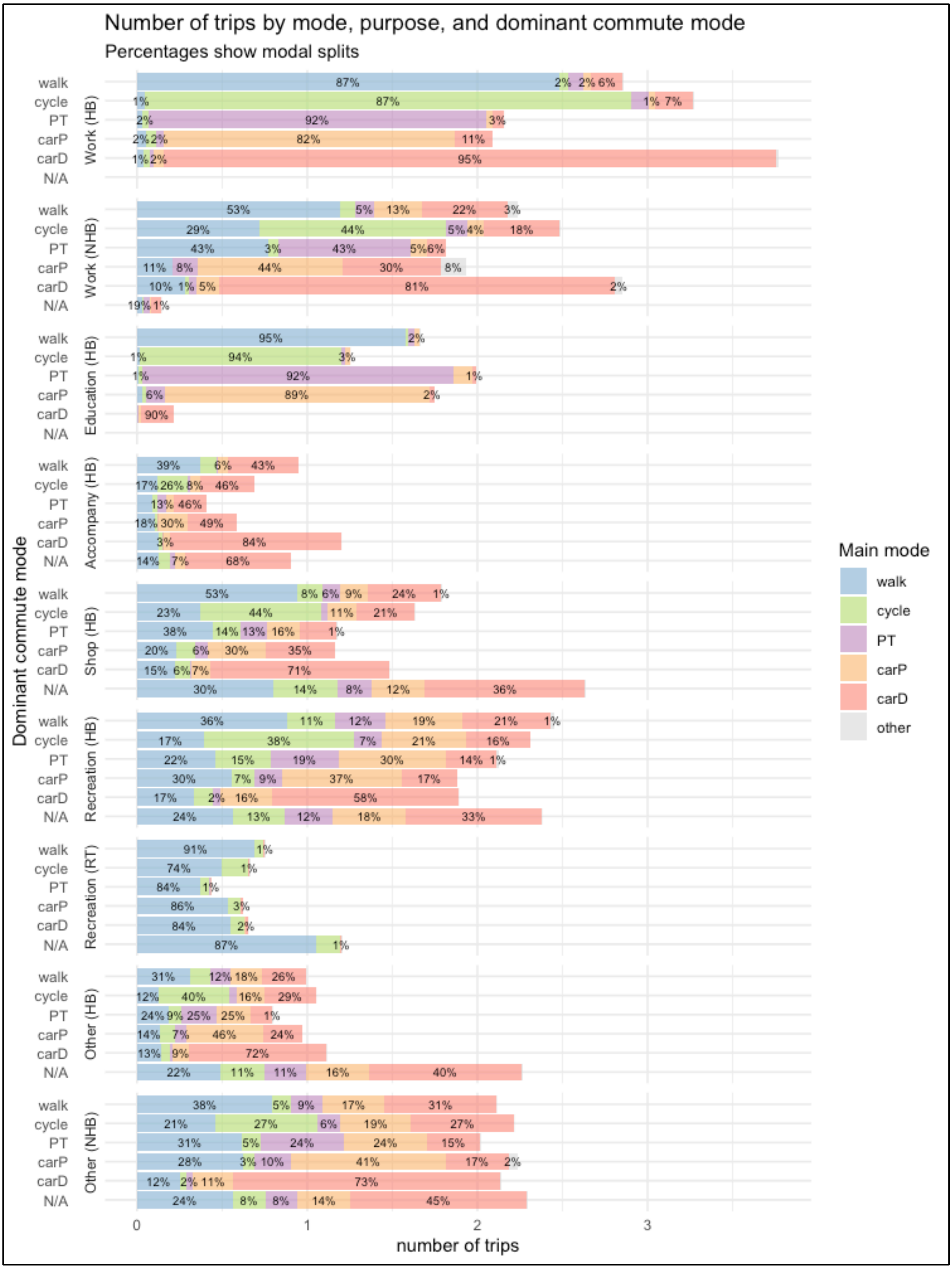


Figure 26: Number of trips by mode, purpose, and dominant commute mode

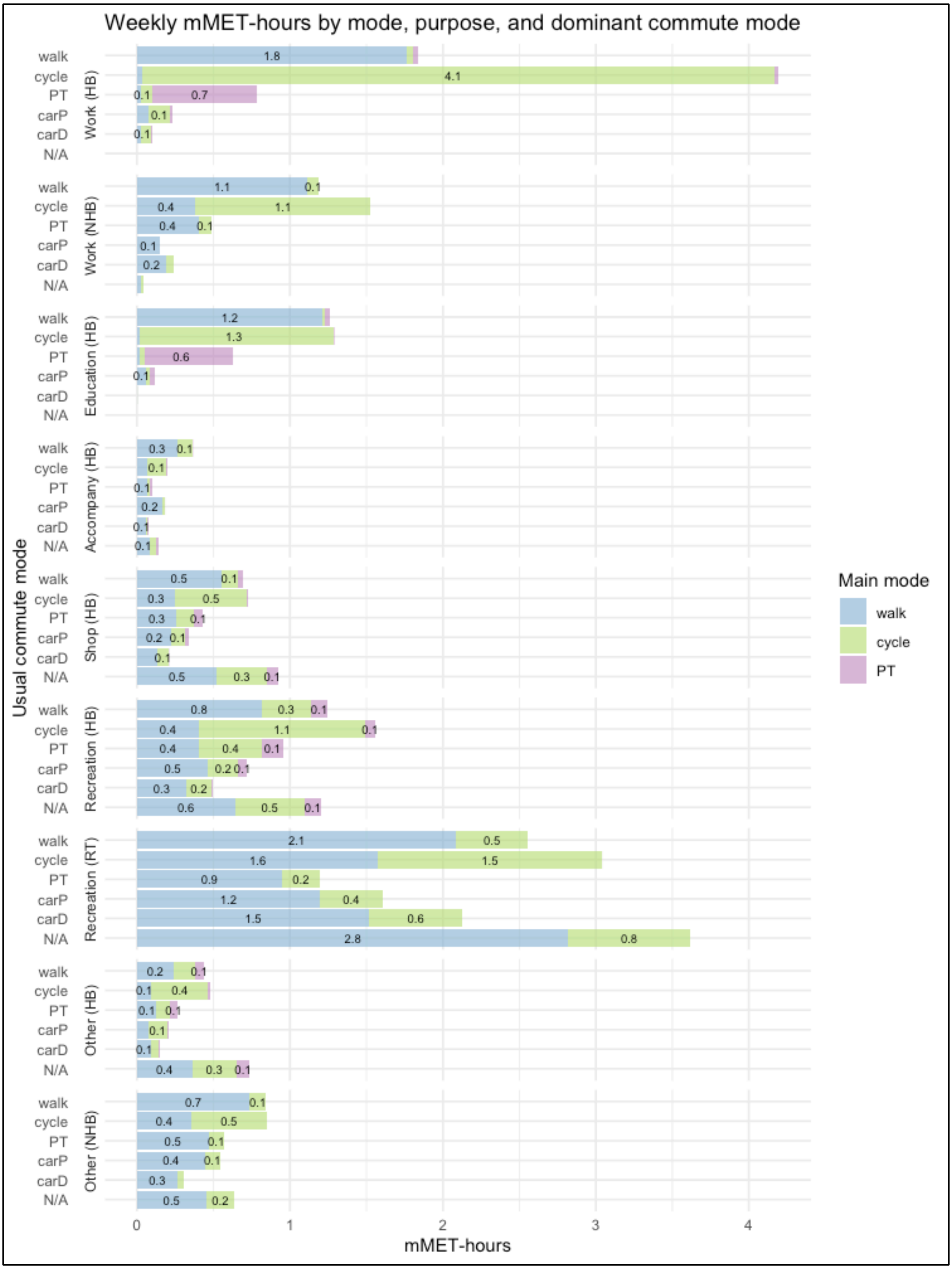


Figure 27: Weekly mMET-hours by mode, purpose, and dominant commute mode

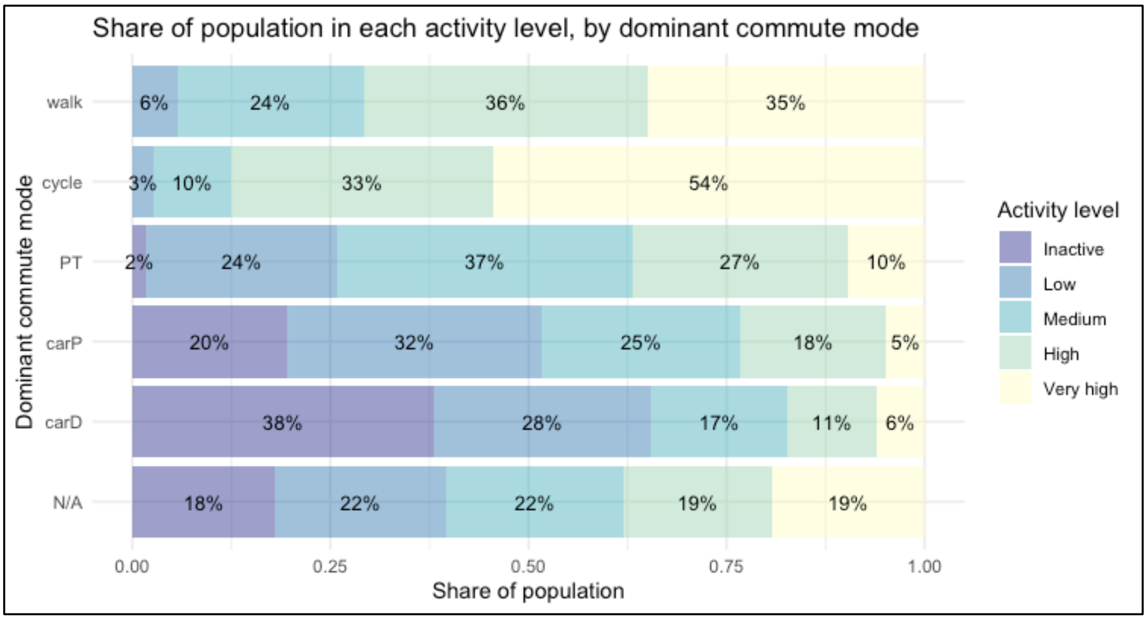


Figure 28: Share of population in each activity level, by dominant commute mode

5.3. Assessment using a Single-day versus a Seven-day Diary

This section investigates the impacts of assessing transport-domain PA using single-day diary data rather than a week-long diary. To perform this analysis, the MOP is compared to an equivalent single-day version of the MOP in which person IDs are not carried over from day to day. The data behind both versions are identical (except for person IDs), enabling a fair comparison. The mean PA volume is the same in both versions, and only the shape of the distribution is different.

For this analysis, PA volumes were converted to RR using the dose-response functions from Kelly et. al. [12] as described in the methodology (section 3.5). Key differences between the 1-day and 7-day versions are given in table 6. Figure 29 compares the shape of the individual-level PA distribution, and figure 30 compares the share of individuals falling into each activity level.

These results support the hypothesis that a 1-day diary is not representative of the weekly distribution of transport-domain PA. The proportion of inactive individuals is overestimated, while the number of individuals in the central PA levels is underestimated. The proportion of individuals in the highest PA level is slightly overestimated.

Table 6: 1-day vs. 7-day diary assessment	7-day diary	1-day diary
% Active	80%	51%
% Walkers	73%	40%
% Cyclists	32%	16%
Properties of mMET distribution		
Mean	9.24	9.24
Standard deviation	11.6	17.0
Coefficient of variation	1.25	1.84
Properties of relative risk (RR) distribution		
Mean RR (linear)	0.894	0.894
Mean RR (log-linear)	0.903	0.910 (+0.78%)
Mean RR (power 0.75)	0.909	0.919 (+1.10%)
Mean RR (power 0.50)	0.918	0.935 (+1.85%)
Mean RR (power 0.375)	0.921	0.940 (+2.06%)
Mean RR (power 0.25)	0.940	0.957 (+1.81%)

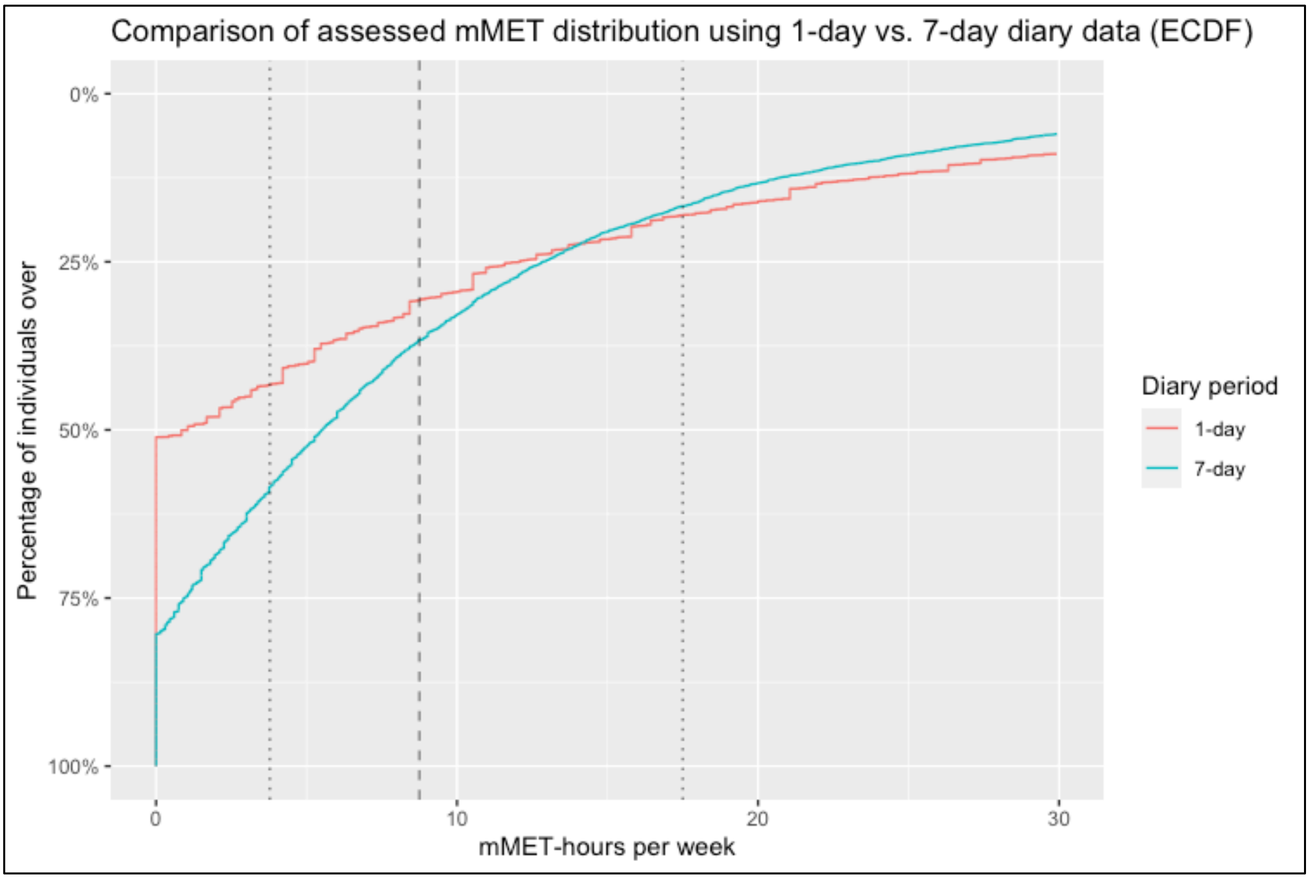


Figure 29: Comparison of weekly mMET distribution using 1-day vs. 7-day diary data (ECDF)

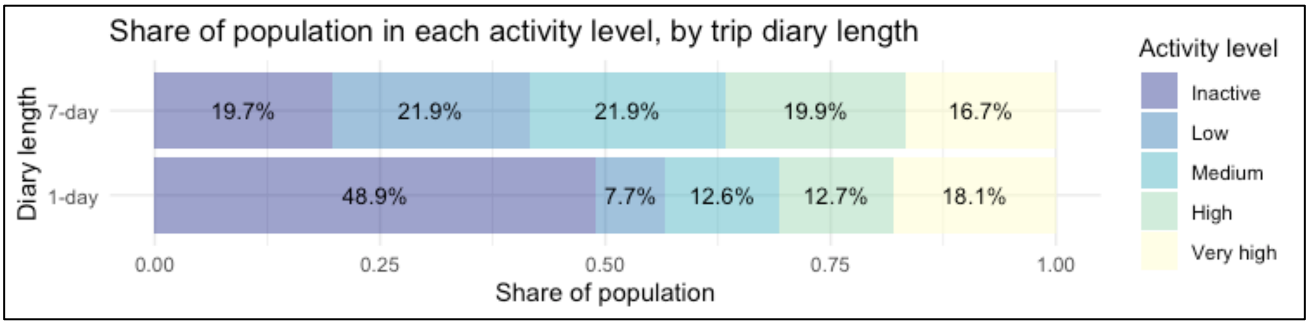


Figure 30: Share of population in each activity level, by trip diary length

Figure 31 compares the baseline day vs. week RR distribution for different dose-response functions. Table 7 shows the mean error in RR for individuals in different activity levels. It shows that mean RR is consistently overestimated for all dose-response curves except linear. There is no error for the linear function because the mean PA volume is the same in both datasets. The RR error is generally highest for the most nonlinear curves. For the steadier curves (e.g. log-linear & power 0.75), the error is larger among high-PA individuals. For the more varying-slope curves (e.g. power 0.5 & 0.25) the error is larger among low-PA individuals.

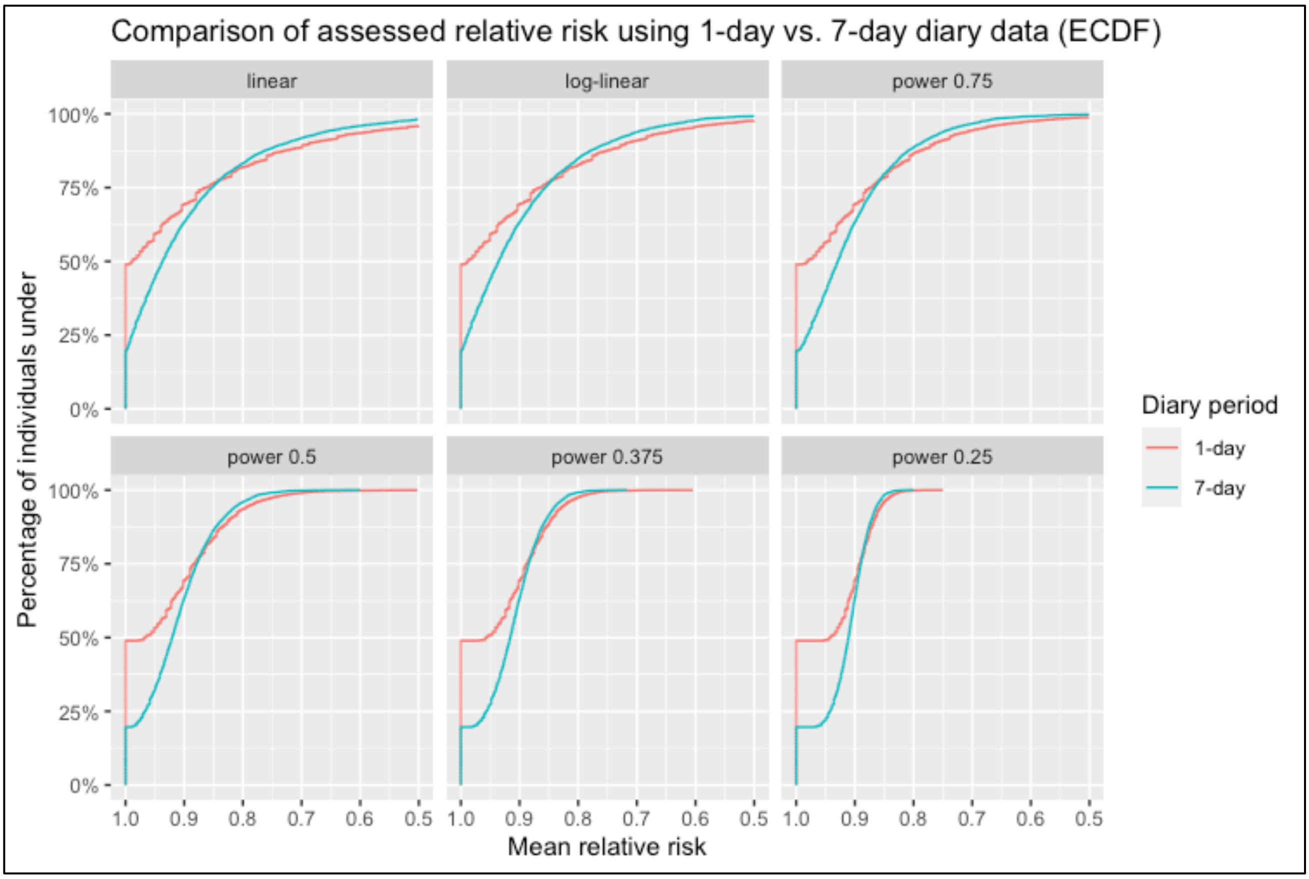


Figure 31: Comparison of assessed relative risk using 1-day vs. 7-day diary data (ECDF)

Table 7: Assessed RR percent error for each activity level and dose-response function

		Physical activity level			
		Low	**Medium**	**High**	**Very high**
Dose-response function	**Linear**	0.00%	0.00%	0.00%	0.00%
	Log-linear	0.09%	0.43%	1.09%	3.61%
	0.75 power	0.79%	1.29%	1.61%	2.20%
	0.50 power	1.97%	2.41%	2.45%	2.43%
	0.375 power	2.79%	2.91%	2.66%	2.26%
	0.25 power	3.85%	3.39%	2.77%	2.00%

Consider a health impact study which investigates the PA benefits of current levels of active travel, versus a hypothetical baseline in which all transport is inactive (similar to Mok et. al. [79]). For such a study, the PAF would be equal to 1 minus the mean RR at baseline. Using the values from table 6, this would cause to error in PAF to range from 7% (using log-linear) all the way up to 28% (using power 0.25). This would correspond to a 7%–28% underestimation in the total number of deaths avoided, a severe error.

Now consider a more typical health impact study in which the current levels are taken as the baseline, and a hypothetical scenario is developed in which transport-domain PA improves in some way. In this case, the extent of the error would depend on how the scenario is defined. For example, if the scenario has a blanket increase in transport-domain PA across the population, there would be an overestimation of health benefits. This is because the baseline RR would be too high, meaning the differential in RR at baseline would also be too high, leading to an overestimation in the PAF. The scale of the error would generally follow the same pattern as in table 7, being larger for curves with a larger change in the differential (like the power 0.25 curve). Now imagine a scenario in which specific groups of individuals (e.g. inactive commuters) are targeted for PA increases. In this case, using a 1-day survey could present major difficulties because it fails to accurately capture inactive and low-activity individuals (see figure 30). This could lead to implausible scenarios in which the number of people targeted are overestimated, which would again lead to an overestimation of health benefits.

Chapter 6. Modelling Trip Generation

This chapter shifts the focus from assessment to modelling. Section 6.1 investigates the suitability of the conventional trip-based paradigm for modelling transport-domain PA, specifically regarding the way trips are coded and the way purposes are defined. Section 6.2 uses correlation plots to explore the relationships between mandatory trip generation, discretionary trip generation, and PA. Finally, section 6.3 develops purpose-specific trip generation models based on the adapted modelling structure defined in chapter 3.

6.1. The Trip-based Paradigm

The trip-based modelling paradigm generally defines six trip purposes: HB work, HB education, HB shop, HB other, NHB work, and NHB other. HB trips are round trips that are modelled in both directions, and NHB trips are modelled in one direction only. To capture HB trips from travel diaries, trips from home (e.g. home–work) are conventionally counted as a round trip, while the return trip to home (e.g. work–home) are neglected. From a travel demand perspective, errors from this simplification process are generally minor and can be addressed with calibration.

The conventional trip-based modelling paradigm leaves an open question about how to handle roundtrips. Following the conventional logic, home-based roundtrips (HBRTs) would be excluded because they end at home, while roundtrips from elsewhere would be coded as NHBW or NHBO trips. However, roundtrips are a major source of PA (recall figure 12), so from a PA epidemiology perspective it is important not to neglect these.

This section explores how the PA distribution would be captured differently under the following paradigms:

(1) Including *all* trips, like in activity-based modelling
(2) Following the trip-based approach, using conventional logic
(3) Following the trip-based approach, including HBRT trips

Both (2) and (3) follow the trip-based approach by neglecting return trips home from elsewhere. However, (3) includes trips that start and end at home, while (2) does not.

Table 8 shows that following the 'conventional' trip-based paradigm significantly underestimates transport-domain PA at baseline. The proportion of walkers is underestimated by about 8%, and the mean distance walked is underestimated by about 29%. The proportion of cyclists is underestimated by abut 5%, and the mean distance cycled is underestimated by about 15%. However, if recreational roundtrips are included in the modelling, this error is nearly eliminated, with the values in option (3) within about 1% of the values in option (1).

A comparison of the mMET distribution at the individual level is given in figure 32. Again, the figure shows that the conventional trip-modelling paradigm (without HBRT) significantly underestimates PA across the population. However, the underestimation is greater for more active individuals, likely because these types of individuals are more likely to take HBRT trips (recall figure 18). Regarding health impacts, this underestimation of baseline PA would lead to an underestimation or overestimation of health benefits, depending on how the scenario is defined. If HBRTs are added to the trip-based paradigm, the distribution is much closer. However, there is still a very slight underestimation towards the bottom of the PA scale, and overestimation towards the top. This is expected, because ignoring return trips home trips takes away some of the variability in the data.

Table 8: Comparison of Trip Based Paradigms

Paradigm	% Active	% Walkers	% Cyclists	Mean km walked	Mean km cycled	Mean mMET hours per week
(1) All Trips	80.3%	73.4%	32.5%	7.32	9.14	9.24
(2) Trip-based (no HBRT)	76.0%	67.8%	30.7%	5.18	7.73	7.28
(3) Trip-based (w/ HBRT)	79.7%	72.5%	32.4%	7.35	9.23	9.35

Figure 33 compares the proportion of individuals falling into each activity group. It shows that neglecting HBRT trips leads to an overestimation of inactive and low-PA individuals, mainly at the expense of individuals at the very-high-PA level. Modifying this to include HBRTs still very slightly overestimates the proportion of individuals at the extreme ends (inactive and very high). However, the overestimation is minor (2–3%) and it can still capture the vast majority of the variability in PA.

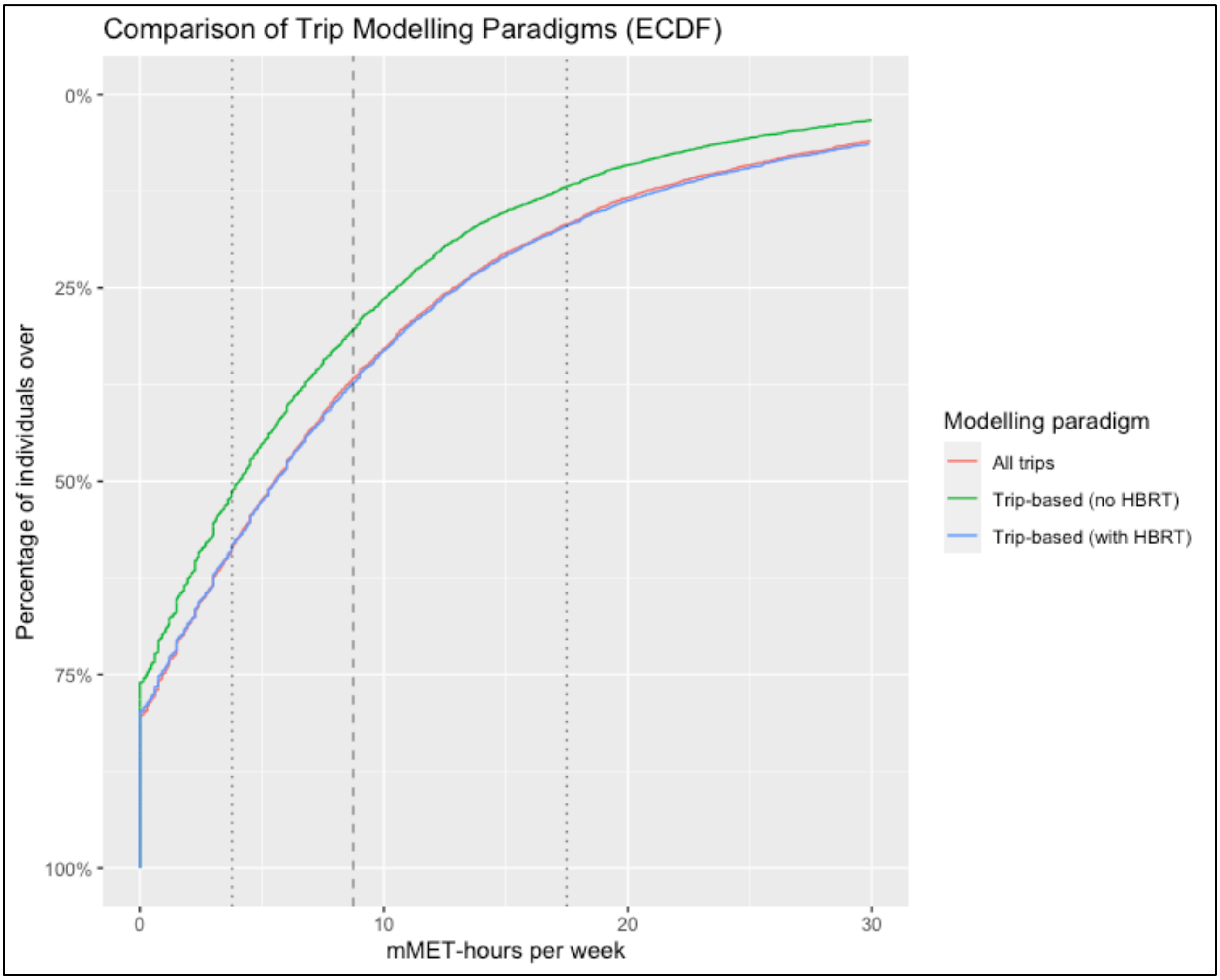


Figure 32: Comparison of Trip Modelling Paradigms (ECDF)

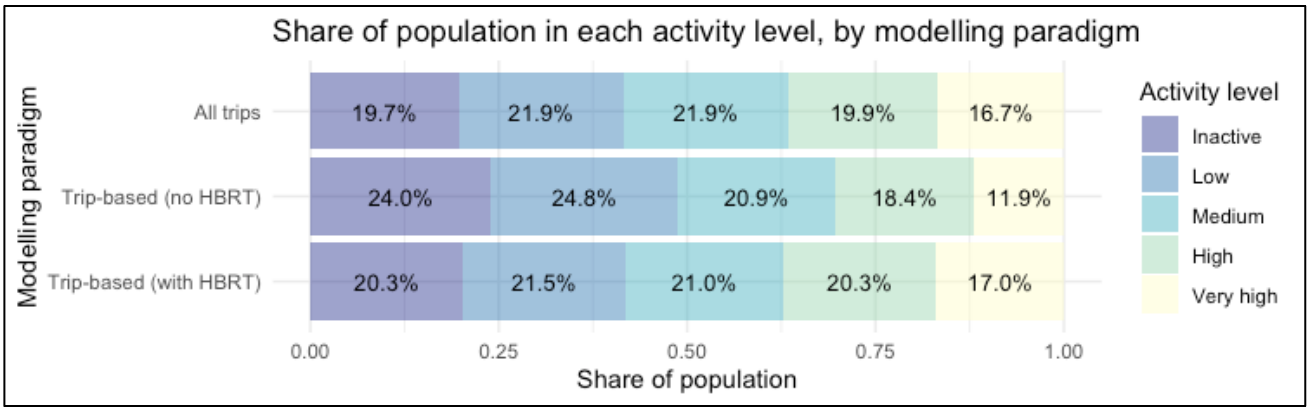


Figure 33: Share of population in each activity level, by modelling paradigm

Based on these results, the remainder of this study uses a modified trip-based paradigm that includes HBRT trips. In the MOP, all HBRT trips are coded as recreational, so these were combined with other recreational roundtrips. Normal (non-round) HB recreation trips were defined as a separate purpose because of the significance of HB recreation trips with regard to PA (see figure 12). HB accompanying trips were not especially relevant for PA so these were combined with HB other trips, which follows convention. The following purposes are defined for modelling:

- Home-based work (HBW)
- Home-based education (HBE)
- Home-based shop (HBS)
- Home-based recreation (HBR)
- Home-based other (HBO)
- Recreational round trip (RRT)
- Non-home-based work (NHBW)
- Non-home-based other (NHBO)

As with conventional trip-based modelling, HB trips are regarded as return trips (and therefore counted twice), NHB trips are one-way (and therefore counted individually), and trips home from elsewhere are not counted. RRT trips were counted individually.

Recall from figure 7 that RRT trips in the MOP are almost entirely (98%) active. Because the sample size of non-active RRT trips is so small, there would insufficient data for reliably estimating mode choice for this purpose for modes other than walking and cycling. Therefore, non-active RRT trips were recoded as either HBR or NHBO trips depending on their origin.

6.2. Correlations in Trip Generation

Correlation plots are used to explore relationship between trip generation for each purpose, as well as the relationship between purpose-specific trip generation and transport-domain PA. To do this, the following was calculated for each individual, using the trip-based (with HBRT) paradigm described in the previous section:

- Whether or not any trips were taken (binary), for each purpose
- The total number of trips taken (count), for each purpose
- The total transport-domain PA

The calculations were performed for the full 7-day diary and for the equivalent 1-day diary (described in section 5.3). Figure 34 shows the correlations for the 7-day diary and figure 35 shows the correlations for the equivalent 1-day diary. The plots on the left show the binary correlation, while the plots on the right show the count correlation. The rightmost columns of each plot show the correlations between trip generation and total individual-level transport-domain PAEE volume (in mMET-hours) for the corresponding time period. These plots were generated using the *corrplot* package in R [80].

These plots reveal an overall negative correlation between mandatory and discretionary trip generation, with the exception of HBW and NHBW which is positive. The mandatory-discretionary relationship is generally stronger for work commutes than for education commutes. The strongest negative mandatory-discretionary relationships are between work trips and HBO (1-day and 7-day) and HBR (1-day only). Correlations with the active-only RRT purpose are weak, suggesting that mandatory trip generation is a poor indicator or how many RRT trips are taken. There is also a negative correlation between NHBW trips and HB discretionary purposes, which suggests that these HB discretionary trips substitute more complex trip chains involving shopping/recreation/other trips that start or end at work.

The relationship between purpose-specific trip generation and PA is analogous to the previous findings in figure 18. The relationship is strongest for RRT trips, which makes sense as it is an active-only purpose. The relationship is also positive for HBS and HBR trips, so taking more of these trips is also an indicator of being more active. However, there is no clear relationship between NHB trips and PA. This suggests that individuals who perform their shopping/recreation/other activities as part of more complex trip chains are less active in the transport-domain than if they had performed them as HB trips. The relationship between mandatory trips and PA is negative but very small. This suggests mandatory trip generation alone is not a strong predictor of transport-domain PA; however, the interactions between mandatory trip generation and other aspects of travel behaviour might still be relevant.

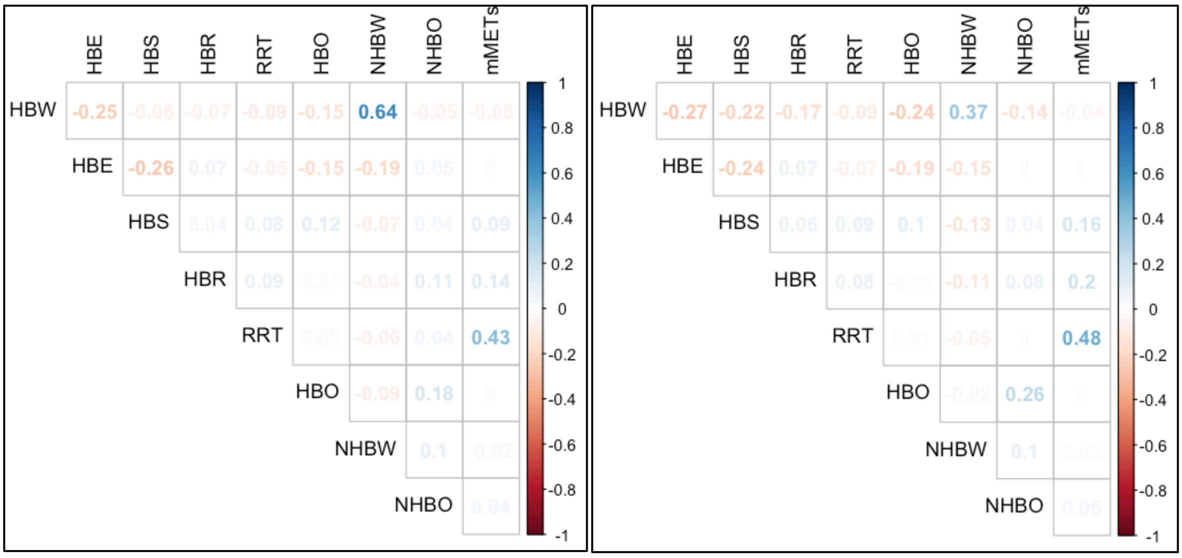


Figure 34: Trip Generation 7-day Correlation Plots, binary (left) and count (right)

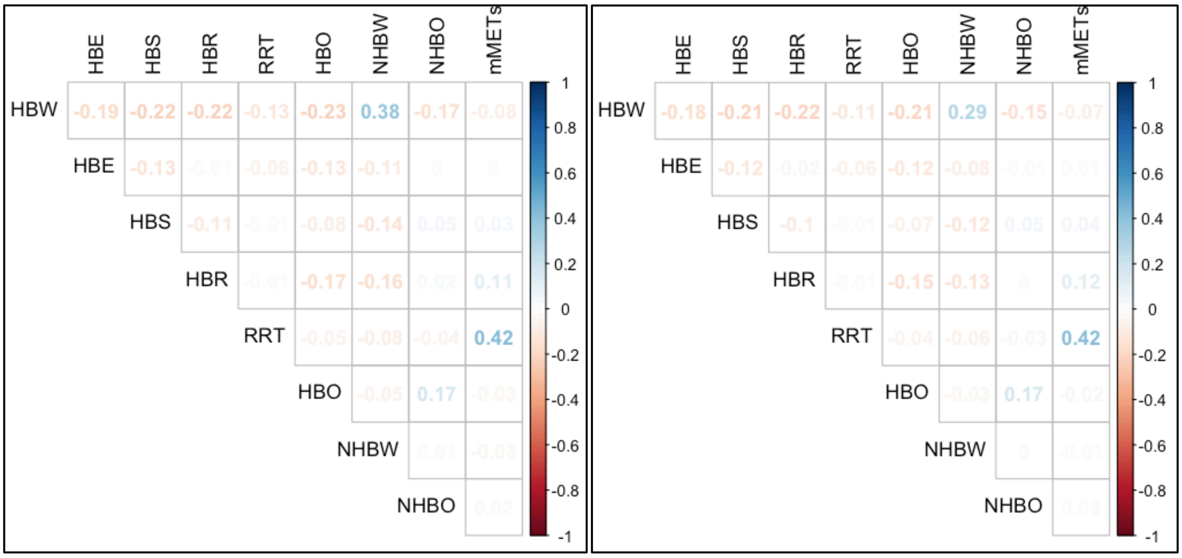


Figure 35: Trip Generation 1-day Correlation Plots, binary (left) and count (right)

Comparing the 1-day and 7-day plots, the relationship between trip generation and total PA is stronger over 7-days, which supports the use of 7-day trip generation data for PA studies. The mandatory-discretionary relationships over 7 days are stronger for HBS, HBO, and NHBW, but weaker for HBR and RRT. This suggests that these HBR and RRT trips, which are the strongest predictors of transport-domain PA, are more likely to be being pushed to non-commuting days (e.g. weekends) but still take place. Therefore, using the single-day correlation alone (e.g. in a single-day model) could underestimate the HBR/RRT trips taken by commuters, thereby underestimating their total transport-domain PA.

6.3. Trip Generation Models

The trip generation modelling structure used for this chapter is given in figure 36. The attributes in green are modelled, whereas the attributes in grey are taken as (constant) inputs.

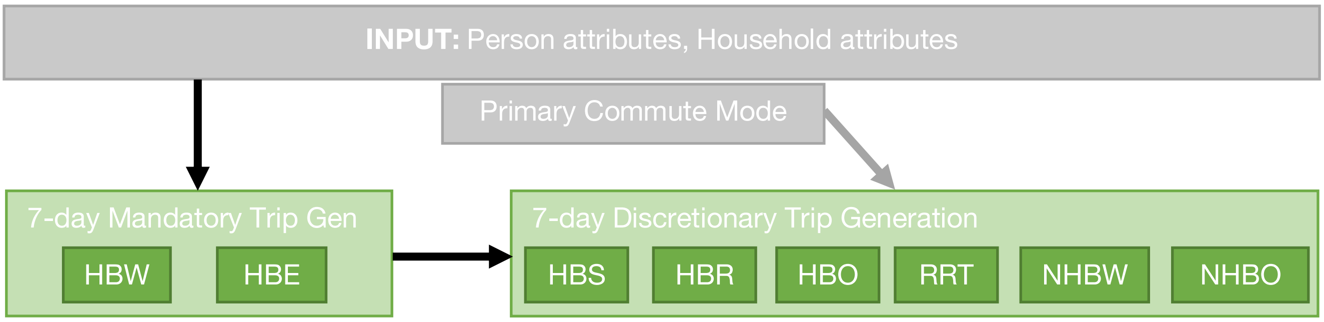


Figure 36: Trip Generation Modelling Structure

The distribution of individual-level trip counts for the 7-day diary are given in figure 37. For comparison, the distribution for the equivalent 1-day diary (as described in section 5.3) is shown in figure 38. This comparison reveals trip generation patterns in the 7-day diary that are not visible in the single-day diary; for example, workers and students are most likely to take five HBW or HBE trips per week, with the second most common number being four. In addition, not all purposes in the 7-day diary are dominated by zeros (the left bar) like they are in the single-day diary. For example, the most common number of HBS and HBR trips taken per week is one, not zero. Nevertheless, several purposes are still zero-heavy in the seven-day diary, including HBW, HBE, RRT, NHBW, and NHBO.

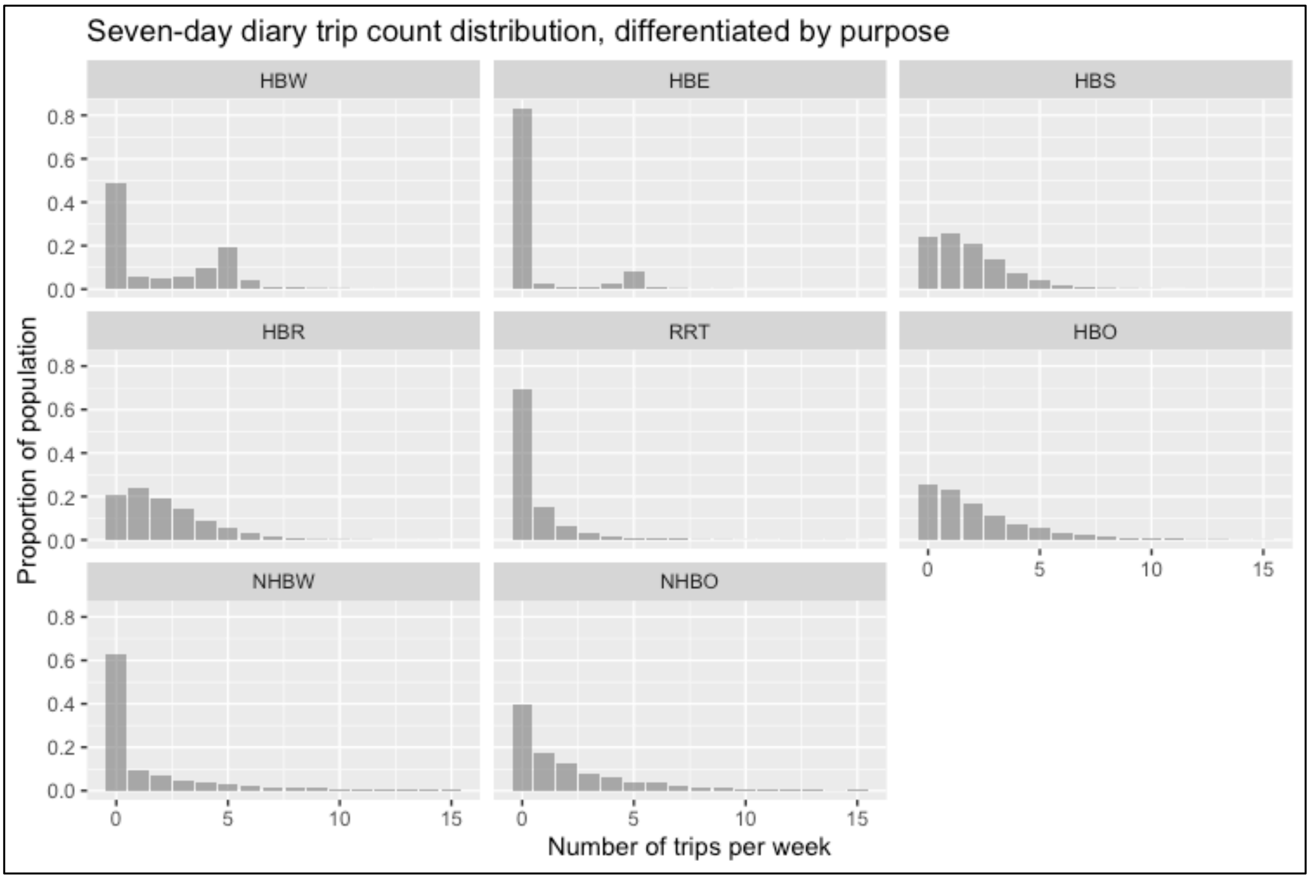


Figure 37: Seven-day diary trip count distribution, differentiated by purpose

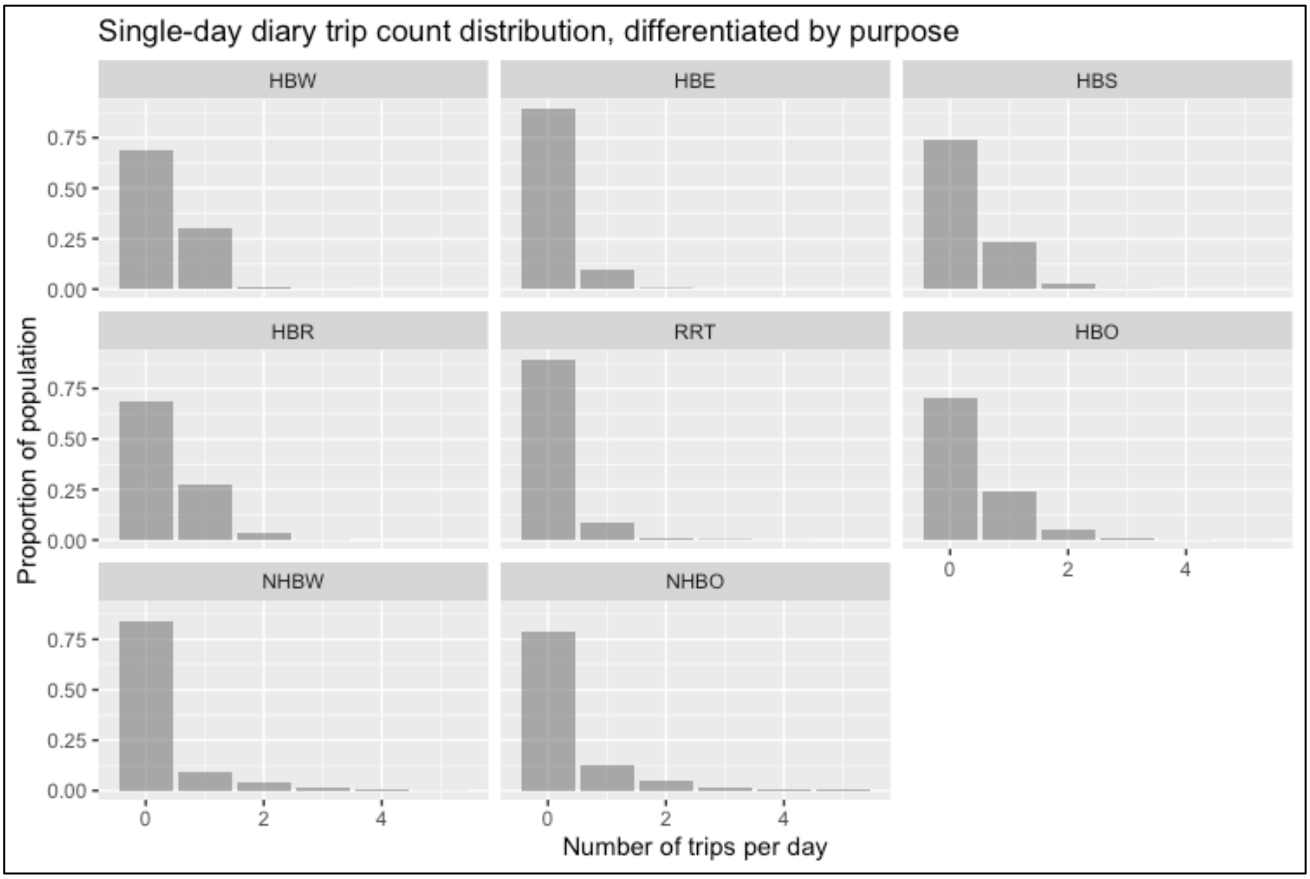


Figure 38: Single-day diary trip count distribution, differentiated by purpose

To avoid modelling an unsuitable distribution due to excess zeros and overdispersion, it was chosen to use hurdle models to estimate 7-day trip generation. With these models, trip generation is modelled in two steps [81]:

1. The ***zero-count*** model estimates whether the individual takes at least one trip (binary logit)
2. The ***normal-count*** model estimates the number of trips taken for those individuals who passed the zero-count step (zero-truncated negative binomial)

The hurdle models were estimated using the *hurdle()* function [82] from the Political Science Computation Library '*pscl*' package [83] in R. Each purpose was estimated separately. To estimate the models for each purpose, a backward elimination stepwise procedure was applied using the independent variables defined in section 3.2. For the discretionary purpose models, the details on mandatory travel behaviour (number of HBW / HBE trips and primary commute mode) were included as additional independent variables, following the modelling structure described in section 3.6.

The estimated model parameters are given in part A of the appendix.

From a PA perspective, of key interest are the determinants of HBR and RRT trips, as these purposes generate the most PA and are the strongest indicators of being active. Cycle ownership increases the odds of taking trips across all purposes (zero-count); however, it was usually not significant for determining the number of trips above one (normal-count). Being elderly is also a significant contributor to RRT and HBR trip generation. Children and young adults are more likely to take HBR trips but less likely to take RRT trips. Interestingly, rural and car-owning households are more likely to take RRT trips, suggesting that RRT trips may be a form of compensation for a lack of other active travel in more car-oriented individuals (a theory previously introduced in section 5.2).

The hurdle model results give further insight into the relationship between mandatory and discretionary trips. In general, more commute trips mean less discretionary trips. For HB discretionary trips, this inverse relationship is largest for HBO trips and smallest for HBR trips, but it is consistently negative. This finding is consistent with the theory of constant travel time budgets. NHBW trips strongly favour commuters, which is clear from the high zero-hurdle coefficients for all primary commute modes. The relationship for NHBO trips is more complex: the results suggest that part-time commuters (1–3 HBW/HBE trips) take more NHBO trips than non-commuters, but full-time commuters (≥5 HBW/HBE trips) take fewer. This is consistent with previous findings in figure 22. Regarding commute distance, longer commutes generally means fewer discretionary trips, which is consistent with the theory of constant travel time budgets.

Furthermore, commuting using active modes (walk & cycle) has a significant impact on trip generation. Active commuters are often significantly more likely to take HB discretionary trips, but less likely to take NHBW trips than car commuters. This supports the theory that taking discretionary trips as HB trips, rather than making them part of more complex NHB trip chains, is an indicator of a more active-transport lifestyle. This is consistent with the findings using correlation plots in the previous section. The relationship between *non-active* commuting and discretionary trip generation remains unclear.

Chapter 7. Modelling Mode Choice

The mode choice model structure for this investigation is shown in figure 39. The full '3-step' model, based on the full model described in section 3.6, is shown using solid black arrows.

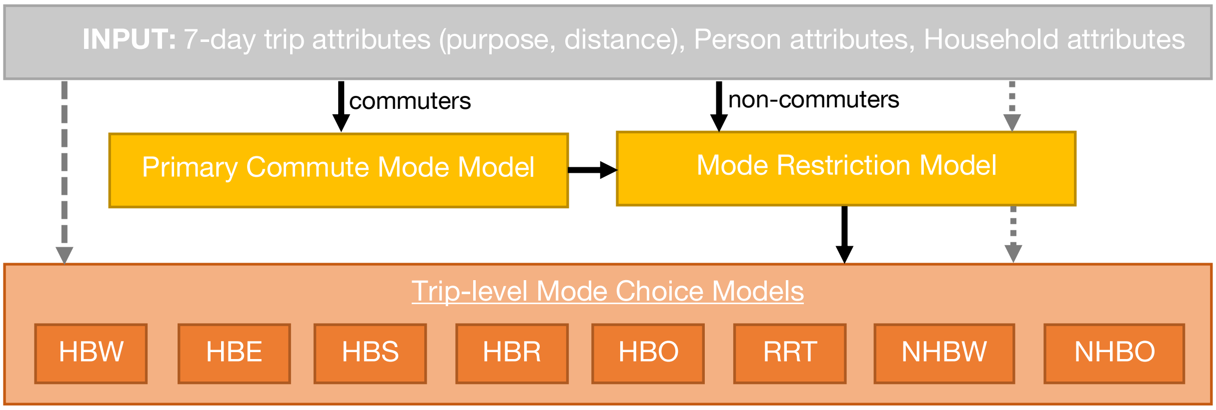
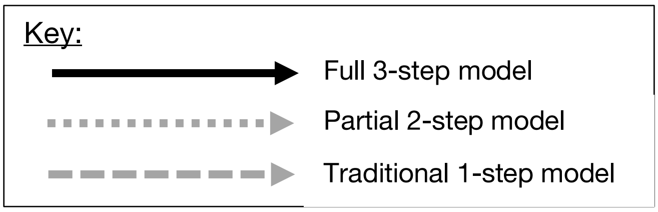


Figure 39: Mode Choice Model Structure



First, a primary commute mode is estimated for anybody who is a commuter. A commuter is defined as somebody who takes at least one HBW or HBE trip during the week (64% of the population). The 'primary commute mode' is defined the same as it was in the assessment phase in section 5.2. Second, the mode restriction model restricts the mode choice set for each individual. Third, the trip-level mode choice models determine the mode for each trip, but the availability of each mode is restricted to the choice set established for each individual in the mode restriction model. Separate mode choice models were estimated for each purpose (8 total).

Recall from section 1.4 that the i.i.d. assumption in logit mode choice models means that *unobserved stability* in individual decisions cannot be captured. This could have important implications for the PA distribution as it may underestimate the proportion of inactive or highly active individuals. The 3-step approach shown in figure 39 aims to overcome this obstacle. Two new individual-level attributes are estimated prior to mode choice, and these attributes become independent variables in the mode choice model. Therefore, more systematic individual-level stability is introduced into the mode choice model, as these previously unobserved attributes (primary commute mode and mode restriction) become observed attributes. In other words, some of the random correlation across decisions in the trip-mode error terms (which cannot be captured) becomes systematic correlation (which can be captured). It is hypothesised that this approach can effectively and plausibly improve the modelled distribution of PA in the population without the need for highly complex modelling structures such as probit or mixed logit.

Sections 7.1–7.4 describe the process for selecting and estimating the parameters of each model. Afterward, section 7.5 evaluates whether this 3-step structure is beneficial for accurately capturing the distribution of PA in the population, or whether steps can be eliminated. For this evaluation, the 'full' 3-step model is compared to an alternative 'partial' 2-step model which has mode restriction but *not* primary commute mode (dotted grey line). Also included in the comparison is the 'traditional' approach in which mode choice is estimated in a single step (dashed grey line). Finally, section 7.6 builds an example scenario to test how health impacts change when nested logit models are used instead of multinomial logit models.

All mode choice estimations and predictions, except for RRT, were made using the *Apollo* package in R [84]. Estimation was performed using a stepwise backward elimination procedure using the independent variables defined in section 3.2. Insignificant coefficients (p-value > 0.1) were eliminated unless they were part of otherwise-significant factor variables or offered meaningful insight into active travel behaviour. Mode choice for RRT was binary (walk or cycle), so this was estimated as a binary logit model in base R. The estimated model parameters for the full (3-step) model are given in part B of the appendix.

This model did not estimate whether PT access and egress was active. Instead, it took a simplifying assumption that PT home-end access was always walked. This is consistent with the assumptions in some previous health impact studies [26, 40]. Because of this, the proportion of active individuals was overestimated by about 1%. However, the PA distribution curve remained virtually indistinguishable, so the effects were assumed to be negligible.

7.1. Selecting a Suitable Nesting Structure

Nested logit models partially relax the i.i.d. assumption by allowing more complex substitution patterns between modes rather than proportional substitution. However, they do not relax the i.i.d assumption across decisions, which remains a limitation as discussed above. For this study, nested logit models were estimated for the dominant commute mode and all trip-level mode choice models except RRT. The nesting structure should be defined based on which modes are assumed to have correlations in their unobserved components of utility. In model estimation, the extent of the correlation can be described by the nesting parameter λ, which should be between 0 and 1 [20]. A value $\lambda = 0$ would indicate perfect correlation, indicating that the choices are essentially the same. A value $\lambda = 1$ would indicate no correlation, which reduces the model to an ordinary logit. A value outside this range is usually inconsistent with utility-maximising behaviour, indicating the model should be re-estimated.

For forecasting, the nesting parameter indicates how much substitution can be expected between modes in the same nest versus modes in other nests. A parameter close to 0 would indicate most substitution comes from within the nest, whereas a parameter close to 1 would indicate substitution is nearly equal. This can be crucial from a PA perspective: as discussed in section 2.3, an increase in cycling would not benefit health if the newly cycled trips substituted for walking trips.

To understand whether there were significant correlations for active modes, three active-mode nesting structures were tested and compared. The first had PT, walk, and cycle in the active nest. The second had only walk and cycle in the active nest. The third had no active nest. In addition to this active nest, the models were initially assumed to have a car nest consisting of car driver and car passenger. While mainly irrelevant from a PA perspective, these car modes are similar and are commonly nested together in mode choice models [53]. However, if the car nest parameter λ_{car} was implausible then the car nest was removed and the active nest models were re-estimated.

The models were initially estimated using all coefficients, with results shown in table 9. The selected nesting structure for each model is shown in bold. Non-significant coefficients were only removed after the nesting structure was fixed for each model (this was to avoid inadvertently introducing more correlation to the error components once variables are removed). While the nesting structure remained fixed, the nesting coefficients were allowed to change (within a suitable range) during the estimation process.

Table 9 reveals that the chosen nesting structure made very little (if any) difference to the overall fit of the model (indicated by McFadden R^2). However, λ_{active} was highly significant in almost all models, which indicates significant correlations in the unobserved portions of utility for active modes. The active nest consistently became even more significant when PT was introduced to the active nest. This is plausible and agrees with the results of previous active travel studies (see section 2.3) that showed higher levels of cycling substitution from walking and PT than from car. Therefore, for all final models it was chosen to include PT in the active nest. Nevertheless, the nesting parameter λ_{active} remains relatively close to one, meaning a substantial proportion of newly active trips would still be drawn from inactive modes.

The car nesting parameter λ_{car} was usually significant and small, indicating close substitution between car driver and car passenger (which is irrelevant for PA). For HBE and NHBW, the car nesting parameter was implausible $(\lambda_{car} > 1)$, so the car nests were eliminated for these purposes.

Table 9: Comparison of Mode Choice Nesting Structures

Model	McFadden R^2	λ_{car} [t-statistic]	λ_{active} [t-statistic]	Active nest structure
Primary Commute Mode	**0.555**	**0.39 [2.24*]**	**0.76 [8.15***]**	**PT + walk + cycle**
	0.555	0.36 [2.11*]	0.19 [0.55]	Walk + cycle
	0.555	0.37 [2.16*]	---	---
HBW Mode Choice	**0.798**	**0.25 [0.75]**	**0.68 [12.38***]**	**PT + walk + cycle**
	0.797	0.18 [0.56]	0.76 [7.73***]	Walk + cycle
	0.797	0.18 [0.56]	---	---
HBE Mode Choice	0.781	1.85 [2.999**]	0.84 [5.446***]	PT + walk + cycle
	0.781	1.96 [3.12**]	1.33 [3.16**]	walk + cycle
	0.781	1.92 [3.11**]	---	---
	0.781	**---**	**0.817 [5.403***]**	**PT + walk + cycle**
	0.781	---	1.29 [3.15**]	walk + cycle
	0.777	---	--	---
HBR Mode Choice	**0.383**	**0.17 [1.731 `]**	**0.62 [20.918 ***]**	**PT + walk + cycle**
	0.381	0.26 [2.26*]	0.86 [19.97***]	walk + cycle
	0.381	0.26 [2.44*]	---	---
HBS Mode Choice	**0.488**	**0.01 [35.2 ***]**	**0.72 [27.15 ***]**	**PT + walk + cycle**
	0.487	0.02 [0.142]	0.85 [19.05 ***]	walk + cycle
	0.487	0.01 [0.128]	---	---
HBO Mode Choice	**0.503**	**0.18 [1.98*]**	**0.70 [22.05***]**	**PT + walk + cycle**
	0.503	0.24 [2.60**]	0.67 [16.69***]	walk + cycle
	0.502	0.25 [2.61**]	---	---
NHBO Mode Choice	0.388	0.31 [2.89**]	0.59 [16.15***]	**PT + walk + cycle**
	0.386	0.46 [4.01***]	0.80 [12.28***]	walk + cycle
	0.386	0.46 [4.02***]	---	---
NHBW Mode Choice	0.587	1.49 [5.14***]	0.72 [17.45***]	PT + walk + cycle
	0.587	1.59 [5.10***]	0.68 [9.56***]	walk + cycle
	0.586	1.66 [5.14***]	---	---
	0.587	**---**	**0.72 [17.52***]**	**PT + walk + cycle**
	0.587	---	0.67 [9.59***]	walk + cycle
	0.586	---	---	---

` ***p-value ≤ 0.1***
* ***p-value ≤ 0.05***
** ***p-value ≤ 0.01***
*** ***p-value ≤ 0.001***

7.2. Primary Commute Mode

The primary commute mode is estimated once for each commuter, and it estimates which mode is dominant in commute trips. This model component does not assign modes to trips. Rather, it adds an attribute to the person which is then included as an independent variable in the mode restriction and mode choice models.

Including a primary commute model aims to achieve two goals. First, it introduces an independent variable to the mode choice models that is consistent across all choices of a given individual. This introduces more systematic correlation to the trip-mode utilities of each individual, thereby adding more stability to individual mode choice behaviour. Second, introducing dominant commute mode as an independent variable in later models can provide further insight into the relationship between mandatory and discretionary travel behaviour. From a forecasting perspective, modelling the primary commute mode first could make it possible to test the overall effects of policy scenarios that specifically target commute behaviour (e.g. a cycle to work scheme).

In general, there are two types of commuters: workers who take HBW trips, and students who take HBE trips. Unfortunately, due to the small MOP sample size and the exclusion of individuals aged ≤10, there was not enough data to estimate this model independently for students. Therefore, only a single model is estimated for both workers and students. The number of HBE trips has become an independent variable in the model, which aims to differentiate students from workers.

The estimated model is provided in table B1 of the appendix. The coefficients appear plausible. Key findings relevant to PA are that commuters in larger households tend to favour car, unless the larger household contains children, in which case non-car modes are favoured. However, the relationship for children may be the result of a correlation between the number of children in the household and whether the individual is a child. Student commuters, indicated by the *HBE trips* attribute, favour PT and active modes. This could partially be because of the strong correlation with age (recall figure 21). Individuals who live in urban areas, and individuals who work or study in locations close to PT stations, are more likely to use PT or active modes to travel to work. Individuals in higher-economic-status households tend to favour PT and Cycle, perhaps because they live and/or work in more dense or affluent areas with superior infrastructure for these modes.

7.3. Mode Restriction

Most individuals only use a subset of all available modes throughout the week. For example, 18% of individuals in the MOP dataset travel exclusively by car, while 13% of individuals recorded zero car trips. As shown previously in table 6, only 80% of individuals recorded active travel during the week, either as the main mode or as PT access. Only 2% of individuals were completely multi-modal (using all available modes at least once). These results support the hypothesis that there is an unobserved stability in the mode choice set each individual is willing to consider. Many agent-based transport demand models like *MITO* [53] and *mobiTopp* [67] disregard this type of stability and assume the full choice set is available for every individual. This assumption could have important implications for PA, which will be explored in section 7.5.

The mode restriction model establishes which modes are available to each individual for their trips. It is estimated once for each individual. This model aims to serve two goals. First, having a consistent restricted choice set for each individual adds further systematic stability to mode choice behaviour. Second, restricting the choice set makes it possible to calibrate the model to differentiate people who do or don't use a particular mode. From a PA perspective, this can ensure that the proportion of active individuals is realistic.

The mode restriction model was a multinomial logit model in which each individual is assigned to one of the following choice sets:

#1: car driver + car passenger
#2: PT + walk
#3: PT + walk + cycle
#4: car driver + car passenger + PT + walk
#5: car driver + car passenger + PT + walk + cycle

For estimation, each individual was placed into the highest set (based on the order shown above) that contained all the modes they used throughout the week. For example, somebody who only drove would be assigned to category #1, somebody who only cycled would be assigned to #3, and somebody who drove and cycled would be assigned to #5. In reality, there were many mode combinations for MOP individuals (35 total). However, this study chose these five because it was the smallest possible set of choice sets that achieves the following goals:

- **Active individuals (#2–5) are differentiated from non-active individuals (#1).** This is the most crucial from a PA perspective. PT is regarded as active in this case because home-end access is assumed to have been walked.
- **Cyclists (#3,5) are differentiated from non-cyclists (#1,2,4).** This is also important from a PA perspective because many health impact studies target cycling specifically. This makes it possible to forecast modal shifts to cycling at the individual level, like in the Woodcock et. al. [45]. An example scenario based on this idea is presented in section 7.6.
- **Car users (#1,4,5) are differentiated from non-car users (#2,3).** This is less relevant from a PA perspective, but it is relevant from a sustainability perspective which is often concerned with car use. This would make it possible to forecast individual-level modal shifts away from private car.

Note that the set of restricted choice sets may itself be restricted for certain individuals. An individual with RRT trips is by definition active, so their choice set options are restricted to #2–5. Similarly, as primary commute mode was estimated in the previous step, the choice set for commuters is restricted to the options that contain this primary commute mode (e.g. if primary commute mode is car, the choices are limited to #1, 4, and 5).

The final mode restriction model coefficients are given in table B2 of the appendix. Note that non-available choices (discussed in the previous paragraph) for certain coefficients are shown in dark grey. The model coefficients show that individuals in urban households and households with children are more likely to travel by active modes at least once during the week. Individuals from higher-income households are more likely to be multi-modal, including car and active modes in their choice sets. Children are also more likely to be multi-modal. As expected, living in a household with many cars is a strong indicator of being car oriented. Females are less likely to include cycling in their choice set. Regarding usual commute mode, PT or walk commuters are much more likely *not* to drive than cycle commuters. Taking more mandatory trips decreased the likelihood of *not* travelling by car but it wasn't a strong indicator of whether somebody was multi-modal. Taking more HB discretionary trips, especially HBR trips, was a strong indicator of being multi-modal. On the other hand, having more NHB discretionary trips indicates a more car-oriented lifestyle. This further supports the theory, previously discussed in section 6.3, that active individuals favour HB discretionary trips while car-oriented individuals favour NHB discretionary trips.

7.4. Mode Choice

Mode choices models for all purposes except RRT were estimated using the nested logit structures given in section 7.1. The RRT models were estimated using a binomial logit model. Because of the 3-step approach, the model estimation process was different from a traditional mode choice estimation in two ways:

The first change is that primary commute mode is introduced as an independent variable in the utility function. Obviously, the primary commute primarily influences commuting trips (HBW & HBE), but it also influences discretionary mode choice. Recall from chapter 5 (figure 26) that individuals tend to favour their primary commute mode for their discretionary trips as well. These models can investigate whether this relationship holds when other independent variables are considered.

The second change is that the mode choice set for each individual is restricted to the choice set established in the mode restriction model. For both estimation and prediction, the probability of choosing any mode outside the restricted set was zero. This was accomplished by assigning choice availability attributes in *Apollo*. Therefore, the mode restriction is also an independent variable in the mode choice model, but it is a hard restriction rather than another utility predictor.

The estimated models are shown in tables B3–B10 of the appendix.

Both mandatory models had high McFadden R^2 values (0.80 HBW and 0.77 HBE), indicating a strong model fit. This was expected to be high since the primary commute mode is an independent variable. Removing this variable would cause a significant drop in the R^2 (to 0.57 HBW and 0.50 HBE). Obviously, the primary commute mode coefficients were highest for their respective mode, indicating that they were in fact primary. The highest primary commute mode coefficients were for PT (in HBW) and walk (in HBE), indicating that these commute modes are highly stable. On the other hand, coefficients for car-passenger-dominant commuters were relatively low, indicating that this mode is more changeable throughout the week. The HBW model included primary commute mode coefficients across all mode options, making it possible to investigate secondary commute modes. For both walk-dominant and cycle-dominant commuters, the second-highest coefficient was PT, indicating that this would be the most favourable alternative if they did not walk or cycle. Another key observation is that the primary commute mode coefficients were *always* positive for modes other than car driver (which was the reference case). This indicates that car driver is indeed the *most* stable mode, whereas those who primarily commute by *any* other mode are more multi-modal. This is consistent with previous findings in figure 26.

The discretionary purposes have more complex relationships with primary commute mode. Across all discretionary purposes, car-driving commuters also favoured car driving for their discretionary trips. This was also the case, albeit weaker, for PT commuters. Car-passenger commuters generally favoured car passenger, but for HBS and NHBO they favoured PT instead. Unfortunately, the relationship between active commute modes and active discretionary trips was weak. For HBS, HBR, HBO, and RRT, the relationship between cycle-commuting and choosing to cycle was either insignificant or negative. They generally preferred car passenger instead. The relationship for walk commuters was also poor or non-existent for these purposes. These poor results for active travellers suggest that including primary commute mode as an independent variable is not an effective way of increasing individual-level active travel stability to mode choices.

Having more frequent commute trips tends to reduce the utility of active modes for mandatory and NHBW trips, but it tends to increase their utility for discretionary trips. For example, commuting full-time decreases the utility of walking to school or cycling to work; however, it increases the utility of cycling for HBS, HBR, and NHBO, and walking for HBO and NHBO.

The utility of active modes decreases sharply with distance, which is expected. However, it increases with distance for PT, and it also increases very slightly with distance for car passenger. For car passenger, this small increase could be plausible because a car passenger journeys require at least two individuals, adding hassle and therefore making it less attractive than car driver for very short trips. However, the increase for PT is debatable. For longer-distance trips it is plausible because intercity PT services can be significantly faster and more comfortable than driving. However, for medium-distance trips that rely on urban/regional PT services this is probably not the case.

Regarding demographic attributes such as age and sex, the results for different purposes are contradictory. For example, being female decreases the utility of cycling for HBW and RRT, but increases it for HBE, HBS, and NHBW. Being elderly decreases the utility of cycling for HBR but increases it for HBS, HBO, RRT, and NHBW. These contradictions could indicate genuine differences between purposes, or they could be a result of sampling bias because the MOP dataset is so small. Ideally, interactions would be introduced to understand these relationships further, but the small data set can make this challenging or in some cases impossible.

7.5. Final Model Comparison

This section evaluates how effectively this model structure captures the PA distribution in the population. The following model structures are compared:

- The "full" 3-step approach described already in sections 7.1–7.4.
- A "partial" 2-step approach that includes the mode restriction model but *not* the primary commute mode model.
- A "traditional" 1-step approach in which mode choice (for each purpose) is estimated in a single step.

For the two alternative model structures, all relevant models and coefficients were re-estimated to no longer include the variables eliminated by the eliminated step(s). The updated coefficients for the alternative models are not shown in this document or the appendix. For each model structure being tested, mode choice was predicted using the estimated probabilities and monte carlo simulation. In the 2-step and 3-step models, a discrete choice was made (using monte carlo simulation) for each internal step before advancing to the next step. The prediction was performed on the same MOP dataset; unfortunately, due to the small dataset size it was not feasible to create separate estimation and validation sets.

To account for random error due to monte carlo simulation, the results shown in this section have been averaged over 10 model runs.

Modal Split

Figure 40 compares the actual mode split (labelled *reference*) to the predicted mode splits. The 'traditional' 1-step model predicted mode split most accurately, while the 'full' 3-step model was the least accurate. Figure 41 shows that the 'full' and 'partial' models underestimated cycling mainly in HBW and HBE and then overestimated it for RRT. Introducing internal steps builds in more potential for error, so having the greatest error in the 3-step model is entirely expected. Nevertheless, the overall predicted mode share remains within 1% of the actual share.

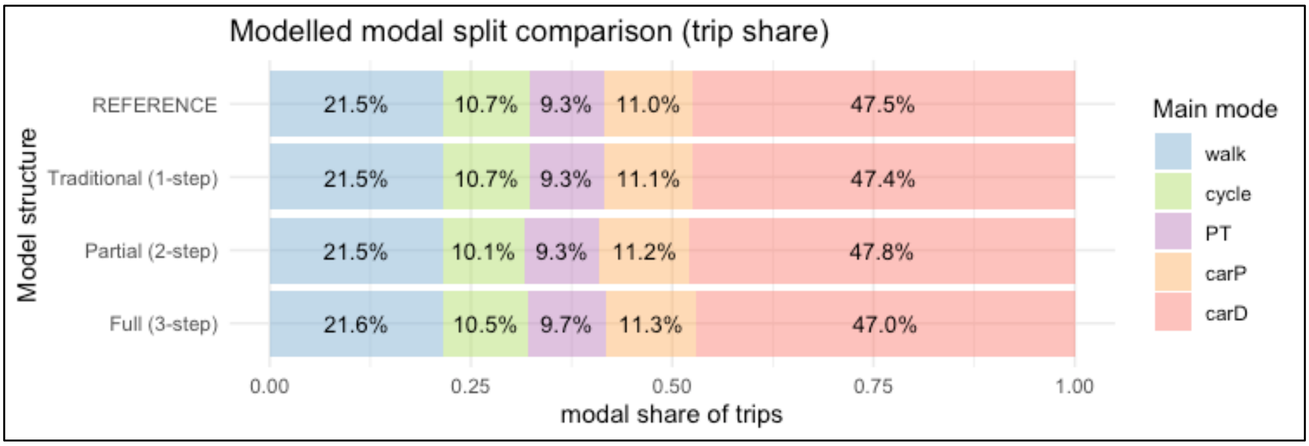


Figure 40: Modelled modal split comparison (trip share)

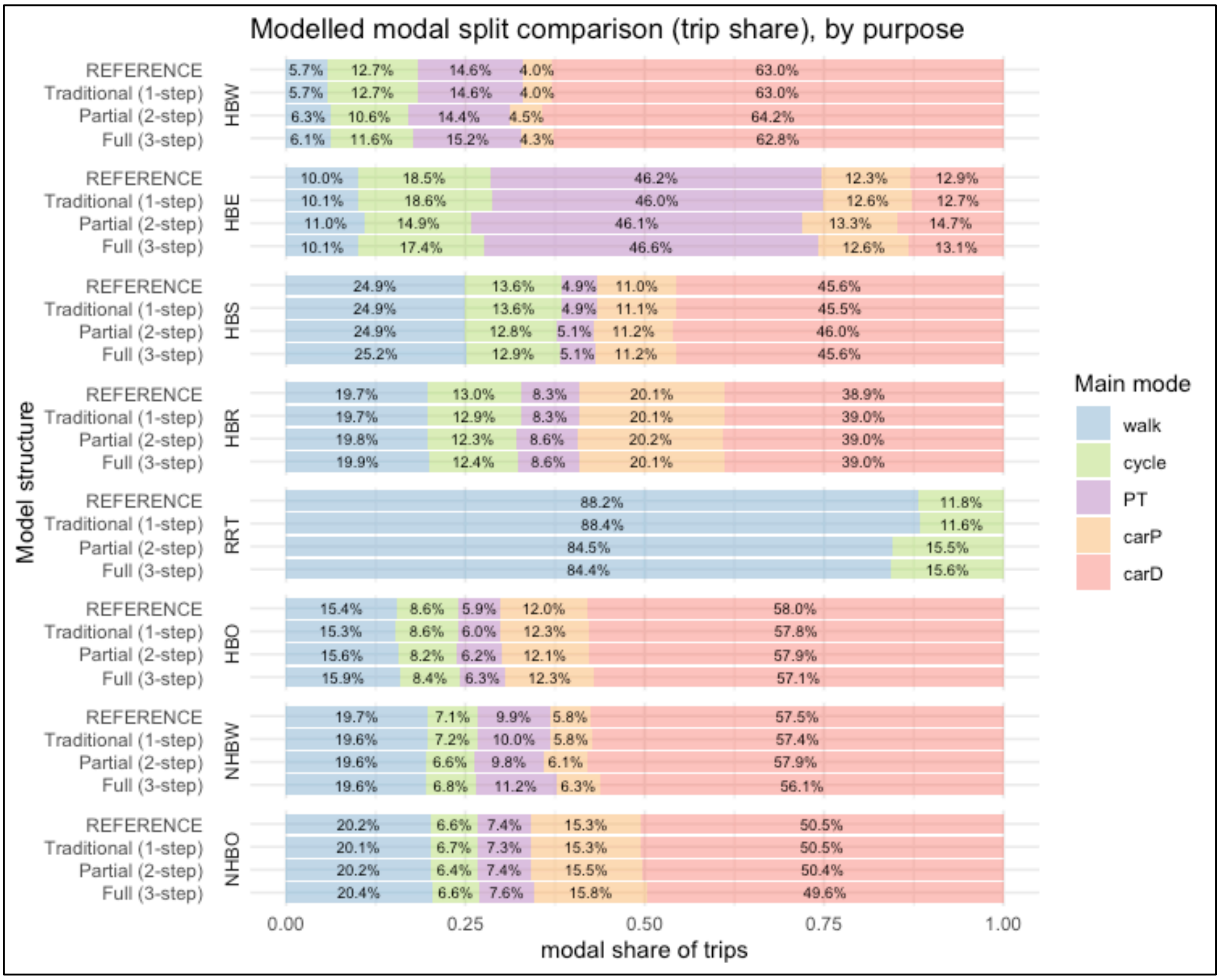


Figure 41: Modelled modal split comparison (trip share), by purpose

Physical Activity Distribution

Figure 42 compares the modelled distribution of PA in the population. Figure 43 compares the proportion of individuals falling into each activity level. For this analysis, the 'traditional' model is applied in two ways. The first, labelled "traditional 1-day", estimates mode choice for a single day and assumes it is representative of the entire week. This is the modelling equivalent to the 1-day assessment shown in section 5.3, and it shows what the modelled PA distribution would look like if a single-day demand model like *MITO* [19] were used for health impact modelling. The second, called "traditional 7-day", estimates mode choice for the full 7-day diary. This shows what the modelled PA distribution would look like if the traditional mode choice modelling approach was applied over a week-long period, like in *mobiTopp* [67].

The results from figures 42 and 43 confirm the hypothesis that traditional mode choice modelling methods poorly capture the distribution of PA in the population. As predicted in section 1.4, the traditional 1-day approach overestimated day-to-day stability while the traditional 7-day approach overestimated day-to-day variability. As a result, the traditional 1-day model overestimated inactive and highly active individuals and underestimated medium-activity individuals, just like the 1-day assessment in section 5.3. In contrast, the traditional 7-day model underestimated inactive individuals, and overestimated medium-activity individuals.

With the ‘partial’ and ‘full’ model structures, the PA distributions were closer to reality. This means the inclusion of the ‘mode restriction’ component was highly effective in introducing systematic stability to individual travel behaviour. The ‘full’ model was only slightly better than the ‘partial’ model, with the distribution curve (figure 42) and activity level share (figure 43) slightly closer to the reference. However, the difference between the two was very small, which suggests that the extra ‘primary commute mode’ component added little value in terms of improving the modelled PA distribution. Unfortunately, both the ‘partial’ and ‘full’ models still significantly underestimated the share of low-activity individuals. This suggests that neither of the new steps were effective for capturing *occasional* active travel behaviour.

As discussed in section 7.3, the mode restriction model made it possible to calibrate the proportion of inactive individuals to the data. Despite this, both the ‘partial’ and ‘full’ models still slightly overestimated the proportion of inactive individuals. This is a limitation of discrete choice methods using monte carlo simulation. The mode restriction model decides which individuals would have active modes available to them. However, just because an active mode is available to somebody does not mean they will definitely choose it for one of their trips. The overestimation of inactive individuals was about 1%, which means about 1% of agents had active modes available but never chose to use them. This issue is also acknowledged in Woodcock et. al.’s ICT [45]. For a well-calibrated model, the mode restriction model could be recalibrated to account for this error.

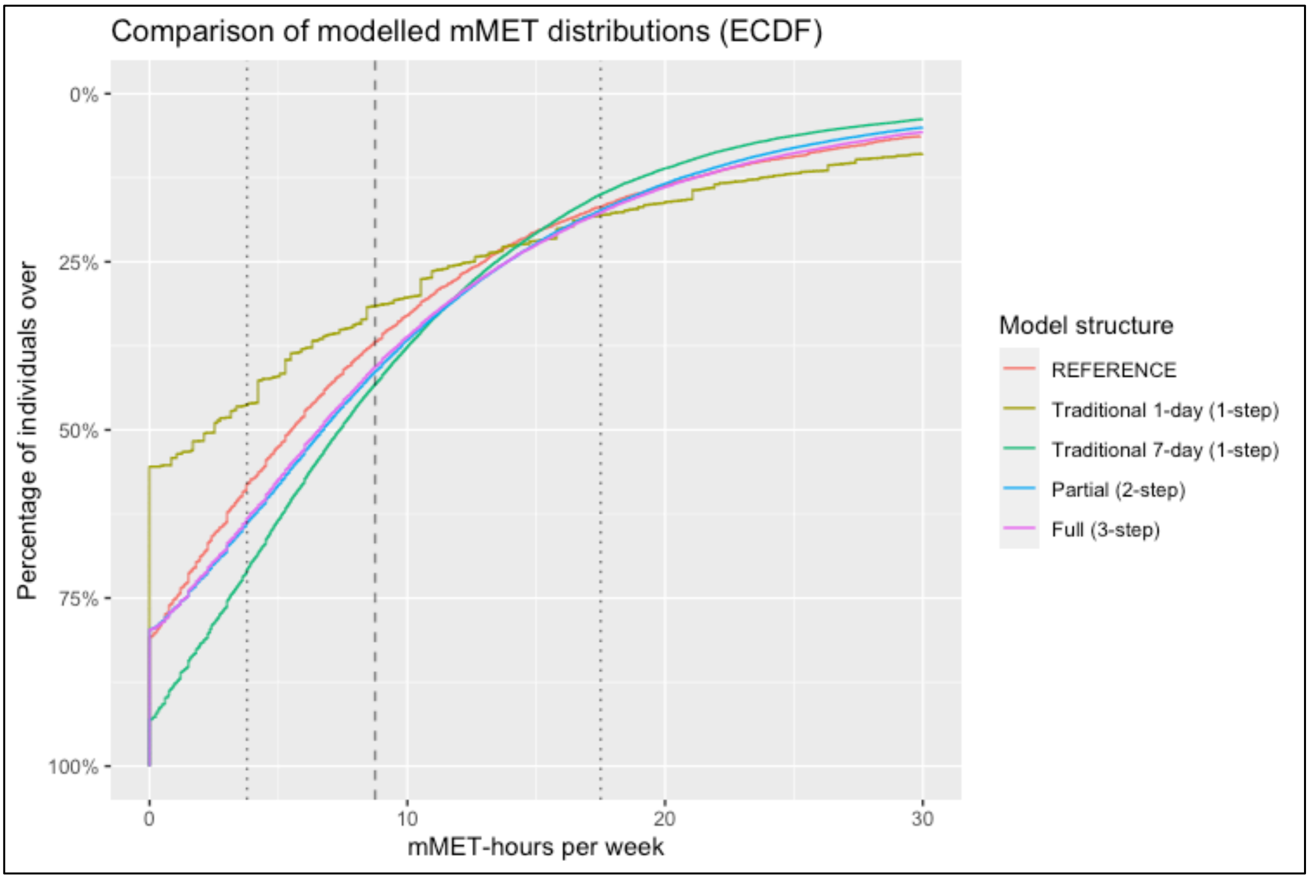


Figure 42: Comparison of modelled mMET distributions (ECDF)

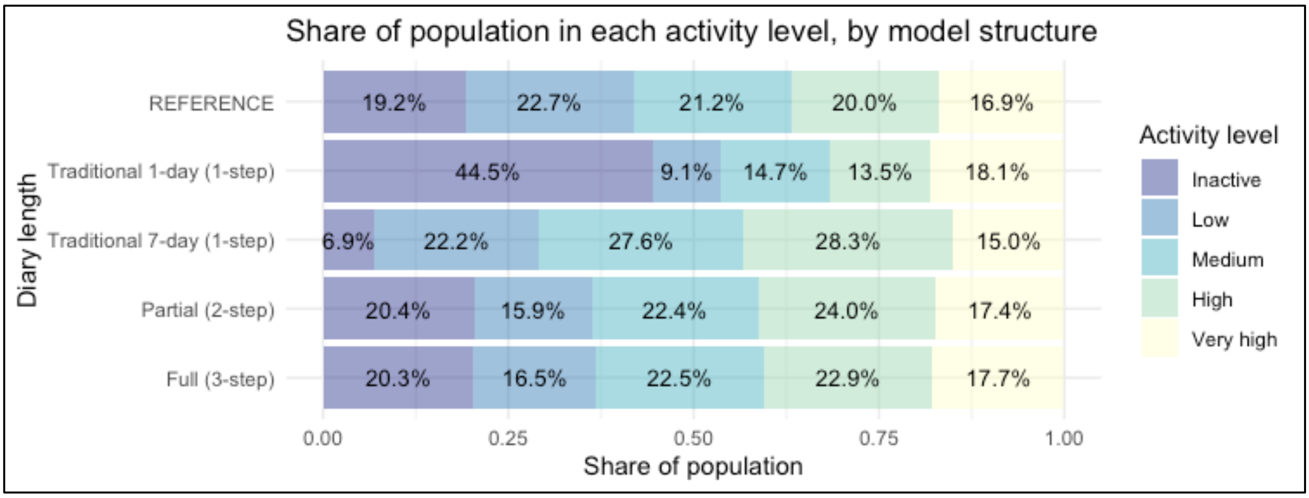


Figure 43: Share of population in each activity level, by model structure

Mode Share Distribution

This section investigates how effectively these models capture the variability and stability in mode-specific choice behaviour. For this investigation, individual-level mode shares were calculated for each agent (e.g. somebody who travelled exclusively by car would show 100% car and 0% for the other modes, or an individual who took half their trips by car and the other half by bicycle would show 50% car, 50% cycle, and 0% for the other modes). Afterward, the individual-level distribution of mode shares is plotted for each mode. These are shown as ECDF plots in figure 44.

The results in figure 44 tell a similar story to the results from the PA distribution shown previously. The 'traditional 1-day' model consistently overestimate the proportion of individuals not using each mode (vertical line at $x = 0\%$), as well as the proportion of individuals *exclusively* using each mode (vertical line at $x = 100\%$). In contrast, the 'traditional 7-day' model consistently underestimated these proportions. For the 'partial' and 'full' models, the accuracy at $x = 0\%$ and $x = 100\%$ was different for each mode. This was entirely expected because of how the mode choice sets were defined for the mode restriction model. Recall the bullet points in section 7.3 describing the goals for defining restricted choice sets. One goal was to differentiate individuals who always travel by car (set #1) as well as individuals who never travel by car (sets #4 and #5); therefore, 0% and 100% car use is well calibrated in the figure. Another goal was to differentiate cyclists from non-cyclists; therefore, 0% cycle use was also well calibrated. A final goal was to differentiate active individuals from non-active individuals; however, this did not differentiate walk from PT (recall that PT trips assume some walking). In other words, the model did not aim to calibrate 0% walkers and 0% PT users separately, but instead just to calibrate 0% for the combination. As a result, the proportion of non-walkers and non-PT users remains underestimated for the "partial" and "full" model. Unfortunately, for this study it was not feasible to calibrate for non-use (and/or exclusive use) of every mode individually. That would require far more than 5 choice sets in the mode restriction model, which is impractical because of the small MOP dataset. It would also have made it more difficult to estimate and interpret the mode restriction model coefficients.

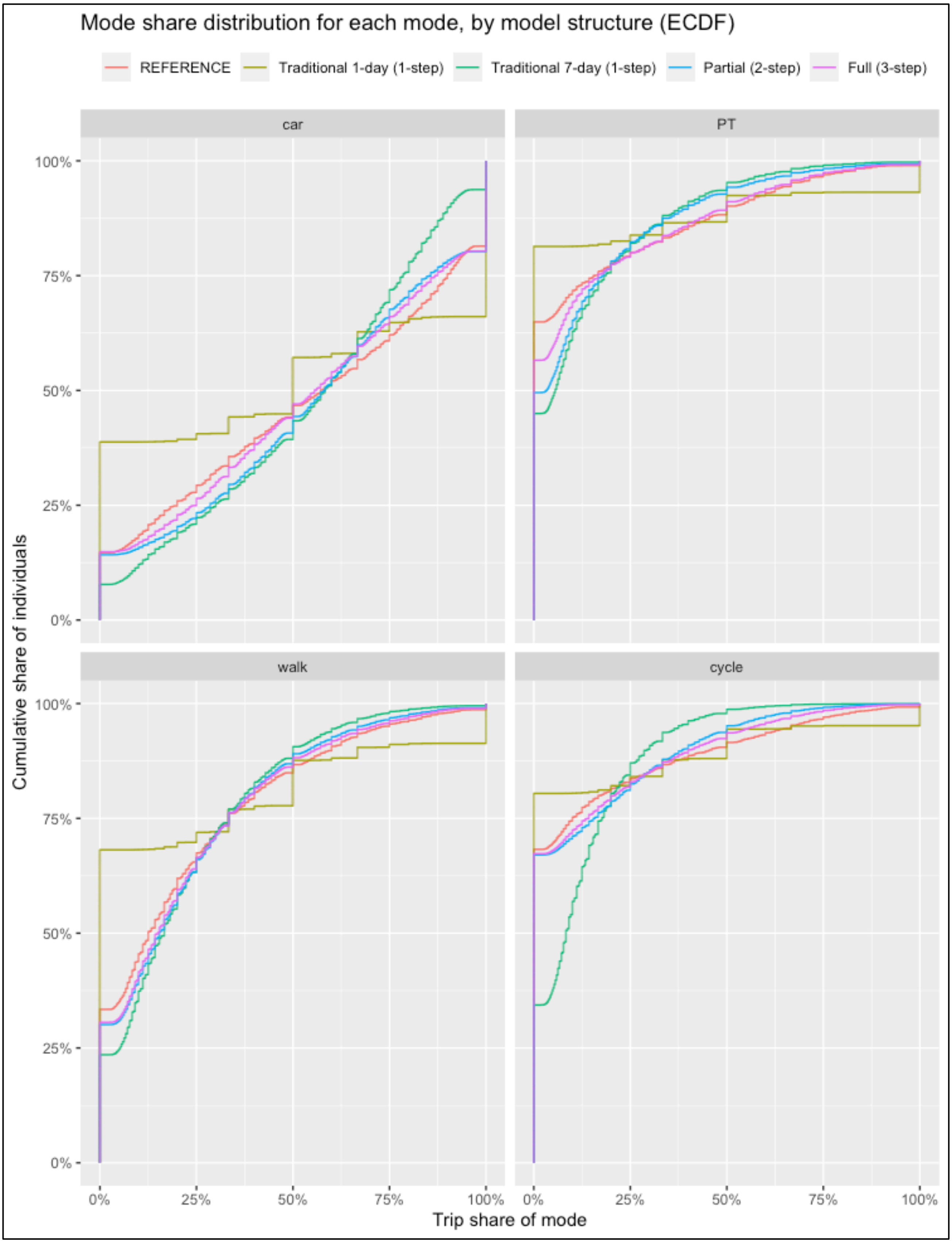


Figure 44: Mode share distribution for each mode, by model structure (ECDF)

For all modes, the 'full' model curve was closest to the reference. The improvement between 'partial' to 'full' was the most noticeable for car and PT. Recall that the difference between the 'partial' and 'full' model was the inclusion of primary commute mode. However, the mode choice model coefficients previously revealed (section 7.4) that the relationship between primary commute mode and discretionary mode choice was strong for car and PT but weak (or non-existent) for active modes. Therefore, it is not surprising that inclusion of the primary commute mode model had the smallest effect on the active mode share distribution.

Relative Risk and Health Impacts

For this analysis, PA volumes were converted to RR using the dose-response functions from Kelly et. al. [12] as described in the methodology (section 3.5). The weighted mean PA volume and RR for each model structure and dose-response curve is shown in table 10.

Surprisingly, the mean PA volume was consistently overestimated by about 5%. This contradicts figure 40, which showed that all model structures were very well calibrated to mode share, actually slightly underestimating active modes. The overestimation of PA volume happens in all model structures, so the source of the error is likely in the trip-level mode choice models. The most likely culprit is the logarithmic transformation used for trip distance in the mode-specific utility functions. While log is a common transformation, it is perhaps unsuitable for active mode utility if the goal is to accurately capture the distance-share of active trips.

Table 10: Modelled mean PA, RR and RR percent error for each model structure and dose-response function

		Model structure				
		REFERENCE	**Traditional 1-day (1-step)**	**Traditional 7-day (1-step)**	**Partial (2-step)**	**Full (3-step)**
Mean PA Volume (mMET-hours per week)		9.31	9.77 (+5.00%)	9.77 (+5.00%)	9.68 (+4.02%)	9.85 (+5.88%)
Mean RR Dose-response function	**Linear**	0.894	0.888 (−0.59%)	0.888 (−0.59%)	0.889 (−0.48%)	0.887 (−0.70%)
	Log-linear	0.902	0.906 (+0.45%)	0.895 (−0.81%)	0.897 (−0.52%)	0.896 (−0.64%)
	0.75 power	0.909	0.915 (+0.66%)	0.900 (−0.93%)	0.905 (−0.45%)	0.904 (−0.54%)
	0.50 power	0.918	0.931 (+1.43%)	0.907 (−1.14%)	0.914 (−0.36%)	0.914 (−0.38%)
	0.375 power	0.920	0.936 (+1.74%)	0.909 (−1.21%)	0.918 (−0.28%)	0.917 (−0.29%)
	0.25 power	0.921	0.940 (+2.05%)	0.910 (−1.26%)	0.920 (−0.18%)	0.920 (−0.18%)

Because the mean PA volume was consistently overestimated, the mean linear-function RR was consistently underestimated. For the non-linear functions, the RR error showed different but predicable patterns based on the function and model structure. The 'traditional 1-day' model behaved similarly to the 1-day assessment analysed in section 5.3: it overestimated RR when nonlinear functions were used, with the most nonlinear curves having the greatest overestimation. All other model structures consistently underestimated RR. The ‘traditional 7-day’ model had the greatest underestimation for all non-linear functions. The ‘partial’ and ‘full’ models had the lowest errors. The errors were smallest for the most nonlinear functions; however, it is difficult to say whether this would still be the case if the mean PA volume hadn’t been overestimated. Had the mean PA volume been properly calibrated, the linear RR error would have been zero, and the more non-linear curves might have slightly overestimated RR. However, the scale of the error would still have been far smaller than the errors for the “traditional” model structures. Based on these RR errors is not conclusive whether the “full” model has any benefits over the “partial” model.

As in section 5.3, the influence of these errors on predicted health outcomes would vary depending on how scenarios are defined. If the scenario compared modelled data to a completely inactive baseline, then the PAF would be equal to 1–RR and the health outcome errors would be severe (up to 24% error for the ‘traditional’ models or 6% error for the ‘full’ model). If the scenario were to define a population-level shift in PA, the errors would probably be smaller and follow the same general pattern as the RR errors in table 10 (i.e. greatest error in the ‘traditional’ models). If the scenario were to target specific groups, such as inactive people or non-cyclists, then using the ‘partial’ or ‘full’ models would be especially advantageous. Because of the mode restriction component, these models are calibrated to properly capture the proportion of individuals in these groups. Using a ‘traditional’ 7-day model would underestimate the proportion of *non*-users, so it would underestimate the individuals changing their behaviour and therefore underestimate health benefits. Using the ‘traditional’ 1-day model would have the opposite effect.

7.6. Nested Logit versus Multinomial Logit

This section defines an example scenario to compare health impacts estimates when using nested logit (NL) mode choice models versus multinomial logit (MNL) mode choice models. The scenario imagines a policy that targets non-cyclists and encourages them to become cyclists. To model the scenario, a change is made to the mode restriction model that doubles the utility of becoming a cyclist. In other words, the utility of choosing choice set #3 (PT + walk + cycle) or choice set #5 (car driver + car passenger + PT + walk + cycle) is doubled for all individuals. The mode choice utilities do not change between baseline and scenario, but a greater proportion of individuals will have cycling available in their mode choice set. This scenario is a simplified version of Woodcock et. al.’s ICT scenarios [45] in which the proportion of people willing to cycle increases. For simplicity, this scenario is based on the ‘partial’ 2-step model structure rather than the ‘full’ 3-step model structure. The baseline and scenario are tested once using NL mode choice models and then again using MNL mode choice models. As in the previous section, the results presented here are averaged over 10 model runs.

The NL nesting structures are defined as in section 7.1, with PT, walk, and cycle sharing a nest. This means new cycle trips will replace a greater proportion of PT and walk trips than car driver and car passenger trips. This is consistent with the results from several intercept surveys (discussed in section 2.3). The MNL model assumes proportional substitution between all modes, which is the assumption currently used in health impact tools like the ICT and PCT [44].

Under the modelled scenarios, the proportion of cyclists nearly doubled, increasing from 33% of individuals at baseline to 63% of individuals in the scenario. The trip mode share of cycling also nearly doubles as shown in figure 45. The NL scenario draws more new cycle trips from walking and PT than the MNL scenario. However, this difference between MNL and NL mode shares appears minor.

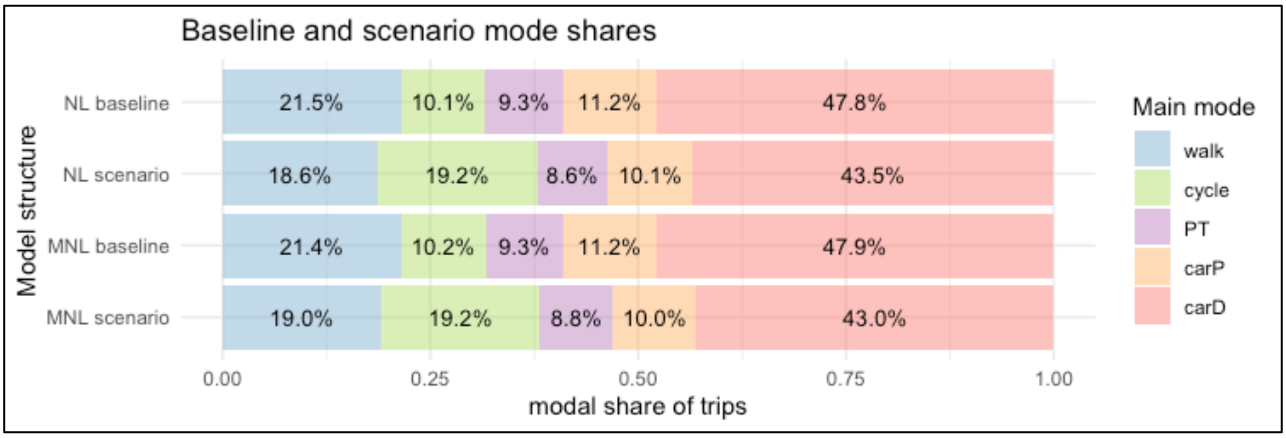


Figure 45: Scenario mode shares

Figure 46 compares the PA volume distributions at each baseline and scenario. As expected, there is a clear increase in PA in both scenarios, but the increase is smaller in the NL scenario where there is more substitution with walking or PT. Specifically, the mean transport-domain PA increases by 2.37 mMET-hours per week in the NL scenario and 2.67 in the MNL scenario. The baseline distributions are similar for MNL and NL, but the scenario distributions diverge. This divergence mainly occurs at higher PA volumes. At lower PA volumes the MNL and NL scenarios are nearly identical. This is because lower-PA individuals would have few or no other active-mode trips to replace, making substitution across active modes less relevant.

A comparison of baseline and scenario RRs and PAFs is given in table 11. In both MNL and NL models, the PAFs are substantially higher for the linear and nearly-linear functions. The linear function PAFs are nearly triple the 0.25-power function PAFs. In other words, the predicted health benefits (i.e. number of deaths avoided) of the scenario are tripled when a linear function is used.

In contrast, the difference in PAF for MNL versus NL is relatively minor. The highest error, at 13.1%, occurs if a linear function is chosen. This means that using an MNL rather than a NL would overestimate the predicted health benefits (i.e. number of deaths avoided) by 13%. This error becomes smaller as the dose-response function becomes more nonlinear. This makes sense because the difference between MNL and NL is mainly at higher PA volumes, where the slopes of the more nonlinear curves are closer to zero and therefore have less impact.

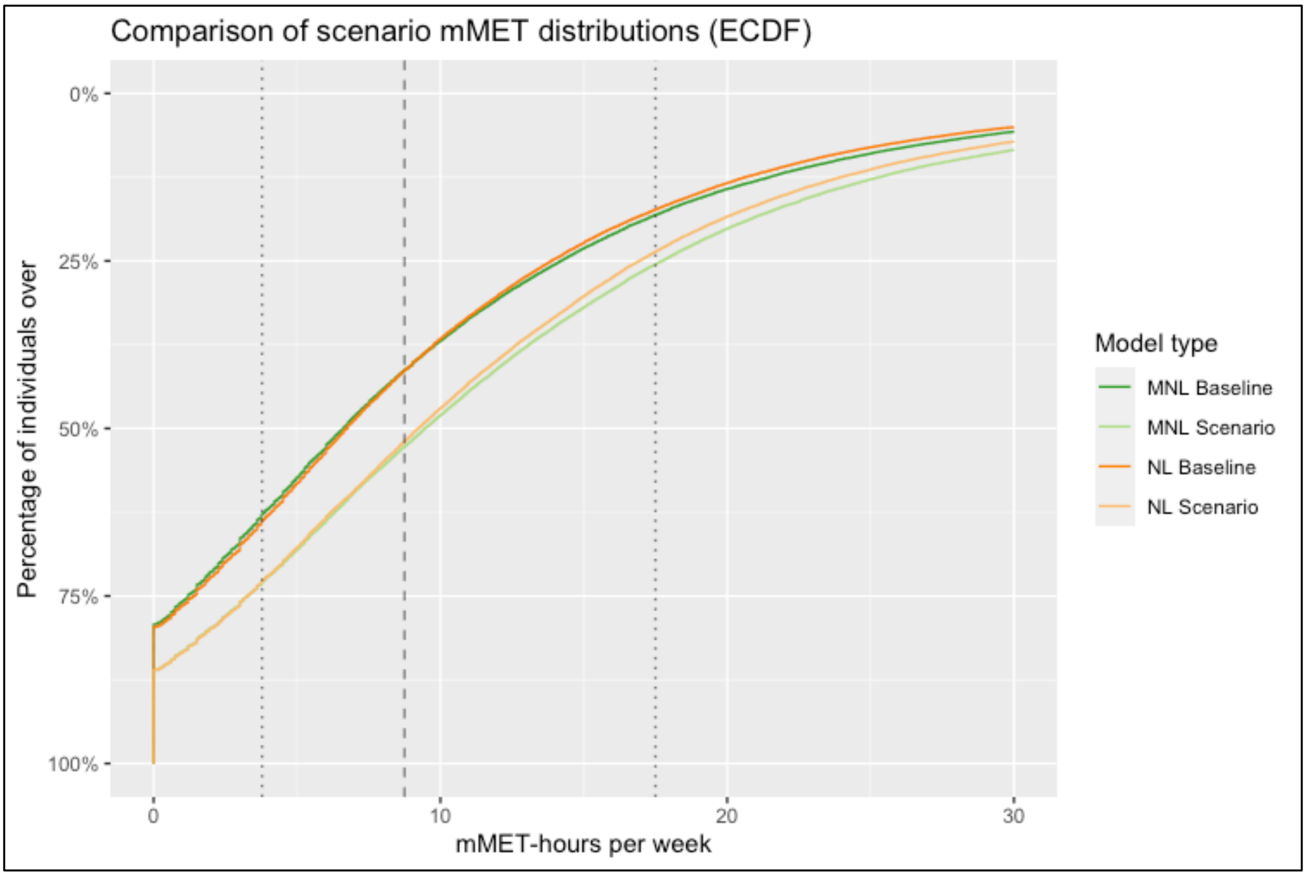


Figure 46: Comparison of scenario mMET distributions

Table 11: Health impact comparison using multinomial vs. nested logit

	Nested Logit			Multinomial Logit			PAF error
	RR baseline	RR scenario	PAF	RR baseline	RR scenario	PAF	
Linear	0.889	0.862	0.030	0.887	0.856	0.034	13.1%
Log-linear	0.897	0.874	0.026	0.896	0.869	0.029	11.8%
0.75 power	0.905	0.885	0.022	0.903	0.881	0.025	11.8%
0.50 power	0.914	0.900	0.016	0.914	0.898	0.017	10.5%
0.375 power	0.918	0.905	0.013	0.918	0.904	0.015	9.9%
0.25 power	0.920	0.910	0.011	0.920	0.909	0.012	9.0%

Chapter 8. Conclusion

The results from this thesis support the central hypothesis that transport assessment and modelling methods would need to be adapted to more accurately evaluate transport-domain PA. The exploratory cross-dataset comparison revealed that to more accurately assess transport-domain PA, HTSs would ideally be adapted to give more attention to:

- Ensuring active RRTs are captured and coded differently from other purposes
- Capturing active trips for the full length of the diary (unlike the British NTS)
- Using stage-level data, to better capture PT access and egress and other multimodal PA
- Avoiding systematic underreporting, which may have been an issue in the British NTS
- Ensuring the data is representative of different age- and sex-specific groups

The following paragraphs discuss key findings regarding the specific adaptations proposed in section 1.5:

Using a 7-day diary instead of a 1-day diary, it is possible to capture more accurately the week-long distribution of PA that is required for health impact tools. This was shown to avoid potentially major errors in health impact estimates; however, the scale of the health impact error would depend on how scenarios are defined. Using a week-long diary also made it possible to identify more complex relationships in travel behaviour that may not have been visible using a single-day diary.

Redefining trip purposes improved the suitability of the trip-based paradigm for modelling PA. First, it ensured HBRT were not neglected, avoiding a potential 27% underestimation of baseline PA volume. Second, it defined new purposes that that were stronger indicators of PA, namely HBR and RRT. This made it possible to more precisely understand and model the determinants and correlates of transport-domain PA.

Mandatory travel behaviour was found to significantly influence discretionary trip generation and mode choice. Therefore, including this in the statistical models provided valuable insight into the way commute behaviour influences discretionary decision making and consequently PA. However, including this relationship in the mode choice models (via the primary commute mode model) only marginally improved the modelled distribution of PA. It did not effectively capture stability in active mode choices, nor did it offer benefits with regard to the health impact indicators. However, only mode choice was tested. It remains to be seen whether including the impact of mandatory on discretionary behaviour would be beneficial if the full model (figure 3) is run. This could be an opportunity for future research

Restricting the mode choice set was highly beneficial for capturing stability in individual mode choice decisions. Including a mode restriction model before estimating mode choice significantly improved the modelled distribution of transport-domain PA, thereby significantly reducing the error in the health impact indicators. However, it was still not effective in capturing *occasional* active travel behaviour, leading to an underestimation of low-PA individuals in the modelled data.

Testing different mode substitution patterns revealed that there were significant correlations in the unobserved portion of utility among cycling, PT, and walking. Via nesting, these correlations were included in the mode choice models, allowing for more plausible substitution patterns that are closer to what has been observed in reality. Neglecting these correlations caused an overestimation in the modelled PA benefits of a shift to active modes. However, this overestimation occured mainly in high-PA groups, so the health impact error is relatively minor.

A key finding from this thesis was the impact of the dose-response function. In the literature review, many previous health impact studies found that the shape of the dose-response relationship significantly influences health impact estimations. This study concludes the same. In fact, the sections of this thesis that calculated health impact indicators (5.3,7.5, and 7.6) found that the sensitivity to the dose-response function was far greater than the sensitivity to the uncertainties being tested. In other words, the uncertainties in assessing and modelling PA that were addressed in this thesis are overshadowed by uncertainties in the health impact estimation process itself.

8.1. Limitations

A major limitation of this study was the limited size and scope of the MOP dataset. For plausible health impact estimations, it is important to clearly differentiate by age and sex. However, the cross-dataset comparison in chapter 4 found that the relationship between PA and age in the MOP is unreliable due to the small sample size. Furthermore, it was generally not possible to control for age and sex in the models (for example, by using interaction variables) because of the limited number of records available. This means many model coefficients were likely influenced by correlations with age, especially those related to students and non-commuters as revealed in figure 21. Information about young children (under age 10) was unavailable in the MOP. This study did not calibrate the PA distributions by age and sex-specific groups, even though this would be an important step in health impact modelling. It is possible that a model built from MOP data would not be suitable for this type of detailed calibration.

The limited size of the MOP dataset was also a bottleneck in many other aspects of model building. For example, because there were so few households it was not feasible to investigate within-household interactions or build household-level trip generation models. The limited size also meant that it was not possible to estimate primary commute mode independently for workers and students, even though these two groups probably behave very differently. The MOP was also not ideal for capturing or modelling PA from PT access and egress, nor was there enough data to differentiate between different types of PT (e.g. train, bus, metro) which can be important in transport modelling.

Another limitation was the way distance was included in the mode choice utility functions. For this study, the distance in kilometres was always transformed using a logarithmic function. However, the distance coefficients for PT and car passenger were sometimes questionable because they were positive rather than negative. Furthermore, the distance share of active modes was consistently overestimated in the models even though the trip share was well-calibrated. Both of these are likely consequences of the logarithmic transformation poorly capturing the relationship between distance and utility. Investigating this relationship more closely could be an opportunity for future research.

While the mode restriction model was effective, the MNL structure may not have been the most suitable. Its coefficients were plausible, but the model fit was relatively poor (McFadden R^2=0.4). The choice sets were defined as to ensure particular groups were well-calibrated (e.g.

non-cyclists, non-active), rather than to match the choice sets commonly identified from the data. This MNL structure would also make it unfeasible to further calibrate other groups (e.g. non-walkers, non PT-users) because this would cause the number of choice sets to increase to an unmanageable amount. Furthermore, choice sets that share a common mode likely have correlations in their unobserved portions of utility; however, these are ignored using MNL. More advanced model structures, such as cross-nested logit or hierarchical binary logit, might be more suitable and/or offer greater flexibility.

This study attempted to introduce systematic stability to mode choice decisions; however, not all forms of stability (or variability) in individual decision-making can be captured systematically. For this study, the estimated mode choice models still assumed i.i.d error terms between choices, so random correlation and variation among decisions for the same individual could not be captured. This is the most likely reason low-PA individuals were still significantly underestimated in the proposed model structure. More advanced model structures, such as probit and mixed logit, may be able to overcome this.

While this i.i.d assumption is limiting in some ways, it adds flexibility in others. This assumption of independence means the trip-specific mode choice models do not need to be trained using longitudinal 7-day data. Instead, they could be trained using larger and more robust single-day datasets such as MiD. Exploiting this i.i.d. assumption therefore presents a major opportunity as it could overcome many of the shortcomings of the small MOP dataset. However, careful attention would need to be given to variable matching between MOP and the alternative dataset, to ensure the mode choice models are consistent with other model components that were estimated using the MOP. Furthermore, individual-level variables would need to be identified in the new dataset that could relate to primary commute mode and the mode restriction. An investigation into the MiD variables found that the mode restriction is available but primary commute mode is not. Therefore, it could still be possible to use MiD data for mode choice if the “partial” 2-step model were used. This is an opportunity for future research.

Another disadvantage of introducing stability systematically, rather than randomly, is that it required building more steps into the model. This increased model complexity and possibilities for random error. It also required additional assumptions about the decision-making process that may not always be plausible. For example, it assumed mandatory trip decisions occurred before discretionary trip decisions. In other words, commuting behaviour is dominant and discretionary travel is built around the commute. However, this might not be the case for everyone, such as part-time workers or individuals with children who might alter their mandatory trips to fit in with their children’s needs.

Finally, there are limitations in the health impact calculation methodology because it has been heavily simplified, as discussed in section 3.5. Future research could introduce more precise PA intensities using available information on gradient, speed, age, and sex. More realistic health impacts would also require PA distributions to be differentiated by age and sex as discussed previously.

8.2. Broader Issues

Open questions remain about the suitability of HTS data for assessing and modelling PA. The cross-dataset analysis in chapter 4 revealed several potential issues in HTSs including underreporting (especially NTS), failure to identify non-mobile individuals (especially MOP), and failure to capture the full scope of transport-domain PA (especially for PT trips). Respondent fatigue is a documented issue, especially in longer-term diaries [71, 85]. Additionally, the findings from the cross-dataset comparison sometimes appeared to contradict results from other studies that used more traditional PA assessment methods. Because of these concerns, the suitability of HTS data for cross-cultural comparisons remains unclear. From a PA epidemiology perspective, this indicates a need to validate HTS assessment against more objective methods, such as those used by Chaix et. al. [46]. As extended GPS and smartphone-based travel diaries become more mainstream [66], there will be more possibilities for future research to validate HTS data against more objective data.

Another open question remains about the suitability of the chosen time period. This study focused on a 7-day period for two reasons. First, travel diaries covering more than 7 days are rare. Second, the 7-day time period is generally consistent with the time period considered for health guidelines, questionnaires, and cohort studies. However, using a 7-day period omits certain types of variation in travel behaviour, such as seasonal variation. For example, a study by El-Assi et. al. [86] found larger seasonal variation in active modes compared to other modes, likely because of the greater influence of weather. Future research could investigate how this longer-term variation might influence health impact results.

8.3. Final Remarks

Without testing the full travel demand model, it is not possible to fully understand how the proposed adaptations affect the accuracy of the modelled PA and health impacts. However, the findings from model building and statistical analyses revealed how these adaptations could be useful for PA policymaking. By including the impacts of mandatory on discretionary travel, policymakers could define new scenarios based on policies targeting commute trips, such as telecommuting or cycle to work schemes. By calibrating for inactivity and non-users of modes (via the mode restriction model), policymakers can more accurately define and target specific behavioural groups in their scenarios. Finally, by establishing new behavioural relationships, using plausible mode substitution patterns, and more accurately capturing the PA distribution, the results from modelled policy scenarios can appear more plausible.

The analyses based on the adaptations put forward also revealed important information about mandatory-discretionary relationships, substitution, and compensatory behaviour that may not have been visible otherwise. Commuters who drove to work were found to consistently and significantly favour car-driving for their discretionary trips, but there was no similar relationship for active commuters. More active individuals tend to perform their discretionary trips as HB trips rather than NHB trips. Work commuters appear to compensate for their lack of work-day recreation trips by taking more on non-workdays. Car-oriented individuals showed a tendency toward taking active RRT trips, possibly as compensation for their car-oriented lifestyle. These findings offer insight into many of the concerns about compensatory behaviour brought forward in previous health impact modelling studies. It also highlights the behavioural complexity of active travel, indicating a need for more advanced model structures to effectively model transport-domain PA.

REFERENCES

[1] J. Lelieveld *et al.*, “Cardiovascular disease burden from ambient air pollution in Europe reassessed using novel hazard ratio functions,” *Eur. Heart J.*, vol. 40, no. 20, pp. 1590–1596, May 2019.

[2] “A European Strategy for Low-Emission Mobility,” *EUR-Lex*, 2016. [Online]. Available: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:52016DC0501. [Accessed: 05-Sep-2020].

[3] World Health Organization (WHO), “Global Status Report on Road Safety 2018,” 2018.

[4] L. C. den Boer and A. Schroten, “Traffic noise reduction in Europe,” 2007.

[5] World Health Organization (WHO), “Climate change and health,” 2018. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/climate-change-and-health. [Accessed: 05-Sep-2020].

[6] R. D. Olson *et al.*, “Physical Activity Guidelines for Americans,” Washington, DC: United States, 2018.

[7] D. S. C. Davies, F. Atherton, M. McBride, and C. Calderwood, “UK Chief Medical Officers’ Physical Activity Guidelines,” 2019.

[8] World Health Organization (WHO), “Global recommendations on physical activity for health,” World Health Organization, Geneva: Switzerland, 2010.

[9] R. Guthold, G. A. Stevens, L. M. Riley, and F. C. Bull, “Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1.9 million participants,” *Lancet Glob. Heal.*, vol. 6, no. 10, pp. e1077–e1086, Oct. 2018.

[10] I.-M. Lee, E. J. Shiroma, F. Lobelo, P. Puska, S. N. Blair, and P. T. Katzmarzyk, “Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy,” *Lancet*, vol. 380, no. 9838, pp. 219–229, Jul. 2012.

[11] L. O’Keefe, “Active Lives Adult Survey November 2018/19 Report,” 2020.

[12] P. Kelly *et al.*, “Systematic review and meta-analysis of reduction in all-cause mortality from walking and cycling and shape of dose response relationship.,” *Int. J. Behav. Nutr. Phys. Act.*, vol. 11, p. 132, Oct. 2014.

[13] S. S. J. *et al.*, “Guide to the Assessment of Physical Activity: Clinical and Research Applications,” *Circulation*, vol. 128, no. 20, pp. 2259–2279, Nov. 2013.

[14] J. R. Speakman, “The history and theory of the doubly labeled water technique.,” *Am. J. Clin. Nutr.*, vol. 68, no. 4, pp. 932S-938S, Oct. 1998.

[15] H. H. J. F. Helmerhorst, S. Brage, J. Warren, H. Besson, and U. Ekelund, “A systematic review of reliability and objective criterion-related validity of physical activity questionnaires,” *Int. J. Behav. Nutr. Phys. Act.*, vol. 9, no. 1, p. 103, 2012.

[16] U. Kunert, J. Kloas, and H. Kuhfeld, “Design Characteristics of National Travel Surveys: International Comparison for 10 Countries,” *Transp. Res. Rec.*, vol. 1804, no. 1, pp. 107–116, Jan. 2002.

[17] A. Ahern *et al.*, “Analysis of national travel surveys in Europe - OPTIMISM WP2: Harmonisation of national travel statistics in Europe,” Jan. 2013.

[18] P. Singleton, J. Totten, J. Orrego-Oñate, R. Schneider, and K. Clifton, “Making Strides: State

of the Practice of Pedestrian Forecasting in Regional Travel Models," *Transp. Res. Rec. J. Transp. Res. Board*, p. 036119811877355, May 2018.

[19] R. Moeckel, N. Kuehnel, C. Llorca, and H. Rayaprolu, "Agent-based Travel Demand Modeling: Agility of an Advanced Disaggregate Trip-Based Model," in *98th Annual Meeting of the Transportation Research Board, Washington, DC*, 2019, p. 20.

[20] K. E. Train, *Discrete Choice Methods with Simulation*, 2nd ed. Cambridge: Cambridge University Press, 2009.

[21] E. Fishman, L. Böcker, and M. Helbich, "Adult active transport in the Netherlands: an analysis of its contribution to physical activity requirements.," *PLoS One*, vol. 10, no. 4, p. e0121871, 2015.

[22] D. Merom, H. P. van der Ploeg, G. Corpuz, and A. E. Bauman, "Public health perspectives on household travel surveys active travel between 1997 and 2007," *Am. J. Prev. Med.*, vol. 39, no. 2, pp. 113–121, Aug. 2010.

[23] J. Pucher, R. Buehler, D. Merom, and A. Bauman, "Walking and Cycling in the United States, 2001-2009: Evidence From the National Household Travel Surveys," *Am. J. Public Health*, vol. 101 Suppl, pp. S310-7, May 2011.

[24] R. Buehler, J. Pucher, D. Merom, and A. Bauman, "Active Travel in Germany and the U.S.: Contributions of Daily Walking and Cycling to Physical Activity," *Am. J. Prev. Med.*, vol. 41, no. 3, pp. 241–250, 2011.

[25] S. Kahlmeier *et al.*, "Health Economic Assessment Tools (HEAT) for walking and for cycling," Copenhagen, Denmark, 2011.

[26] D. Rojas-Rueda, A. de Nazelle, O. Teixidó, and M. J. Nieuwenhuijsen, "Replacing car trips by increasing bike and public transport in the greater Barcelona metropolitan area: A health impact assessment study," *Environ. Int.*, vol. 49, pp. 100–109, 2012.

[27] M. L. Grabow, S. N. Spak, T. Holloway, B. Stone, A. C. Mednick, and J. A. Patz, "Air quality and exercise-related health benefits from reduced car travel in the midwestern United States.," *Environ. Health Perspect.*, vol. 120, no. 1, pp. 68–76, Jan. 2012.

[28] "WebTAG: TAG UNIT A4.1 - Social Impact Appraisal," 2014.

[29] J. Woodcock *et al.*, "Public health benefits of strategies to reduce greenhouse-gas emissions: urban land transport," *Lancet*, vol. 374, no. 9705, pp. 1930–1943, Dec. 2009.

[30] N. Maizlish, N. J. Linesch, and J. Woodcock, "Health and greenhouse gas mitigation benefits of ambitious expansion of cycling, walking, and transit in California," *J. Transp. Heal.*, vol. 6, pp. 490–500, 2017.

[31] N. Maizlish, J. Woodcock, S. Co, B. Ostro, A. Fanai, and D. Fairley, "Health cobenefits and transportation-related reductions in greenhouse gas emissions in the San Francisco Bay area," *Am. J. Public Health*, vol. 103, no. 4, pp. 703–709, Apr. 2013.

[32] G. P. Whitfield, L. A. Meehan, N. Maizlish, and A. M. Wendel, "The Integrated Transport and Health Impact Modeling Tool in Nashville, Tennessee, USA: Implementation Steps and Lessons Learned," *J. Transp. Heal.*, vol. 5, pp. 172–181, Jun. 2017.

[33] W. Nicholas, I. Vidyanti, E. Caesar, and N. Maizlish, "Routine Assessment of Health Impacts of Local Transportation Plans: A Case Study From the City of Los Angeles," *Am. J. Public Health*, vol. 109, no. 3, pp. 490–496, Mar. 2019.

[34] J. Woodcock, M. Givoni, and A. S. Morgan, "Health Impact Modelling of Active Travel Visions for England and Wales Using an Integrated Transport and Health Impact Modelling

Tool (ITHIM)," *PLoS One*, vol. 8, no. 1, p. e51462, Jan. 2013.

[35] T. H. de Sá *et al.*, "Health impact modelling of different travel patterns on physical activity, air pollution and road injuries for São Paulo, Brazil," *Environ. Int.*, vol. 108, pp. 22–31, Nov. 2017.

[36] J. Johan de Hartog, H. Boogaard, H. Nijland, and G. Hoek, "Do the health benefits of cycling outweigh the risks?," *Environ. Health Perspect.*, vol. 118, no. 8, pp. 1109–1116, Aug. 2010.

[37] T. Xia, M. Nitschke, Y. Zhang, P. Shah, S. Crabb, and A. Hansen, "Traffic-related air pollution and health co-benefits of alternative transport in Adelaide, South Australia," *Environ. Int.*, vol. 74, pp. 281–290, 2015.

[38] M. Tainio *et al.*, "Can air pollution negate the health benefits of cycling and walking?," *Prev. Med. (Baltim).*, vol. 87, pp. 233–236, 2016.

[39] N. Mueller *et al.*, "Health impact assessment of active transportation: A systematic review," *Prev. Med. (Baltim).*, vol. 76, pp. 103–114, 2015.

[40] D. Rojas-Rueda, A. de Nazelle, M. Tainio, and M. J. Nieuwenhuijsen, "The health risks and benefits of cycling in urban environments compared with car use: health impact assessment study," *BMJ*, vol. 343, p. d4521, Aug. 2011.

[41] J. Woodcock, M. Tainio, J. Cheshire, O. O\textquoterightBrien, and A. Goodman, "Health effects of the London bicycle sharing system: health impact modelling study," *BMJ*, vol. 348, 2014.

[42] S. Ahmad, A. Goodman, F. Creutzig, J. Woodcock, and M. Tainio, "A comparison of the health and environmental impacts of increasing urban density against increasing propensity to walk and cycle in Nashville, USA," *Cities Heal.*, vol. 4, no. 1, pp. 55–65, Jan. 2020.

[43] R. Lovelace, A. Goodman, R. Aldred, N. Berkoff, A. Abbas, and J. Woodcock, "The Propensity to Cycle Tool: An open source online system for sustainable transport planning," *J. Transp. Land Use*, vol. 10, no. 1 SE-, Jan. 2017.

[44] "User Manual C1: PCT methods for the commuting layer," 2020. [Online]. Available: https://www.pct.bike/manual.html.

[45] J. Woodcock *et al.*, "Development of the Impacts of Cycling Tool (ICT): A modelling study and web tool for evaluating health and environmental impacts of cycling uptake," *PLOS Med.*, vol. 15, no. 7, p. e1002622, Jul. 2018.

[46] B. Chaix *et al.*, "Combining sensor tracking with a GPS-based mobility survey to better measure physical activity in trips: public transport generates walking," *Int. J. Behav. Nutr. Phys. Act.*, vol. 16, no. 1, p. 84, 2019.

[47] J. Panter, E. Heinen, R. Mackett, and D. Ogilvie, "Impact of New Transport Infrastructure on Walking, Cycling, and Physical Activity," *Am. J. Prev. Med.*, vol. 50, no. 2, pp. e45–e53, 2016.

[48] A. T. Moreno and Moe, "Microscopic Destination Choice: Incorporating Travel Time Budgets as Constraints," in *Transport Research Procedia*, 2017.

[49] C. Raux, T. Ma, I. Joly, V. Kaufmann, E. Cornelis, and N. Ovtracht, *Travel and Activity Time Allocation: an Empirical Comparison Between Eight Cities in Europe*. 2009.

[50] D. P. Piatkowski, K. J. Krizek, and S. L. Handy, "Accounting for the short term substitution effects of walking and cycling in sustainable transportation," *Travel Behav. Soc.*, vol. 2, no. 1, pp. 32–41, 2015.

[51] J. de D. Ortúzar Salar, *Modelling Transport*, Fourth edi. Chichester, West Sussex, United Kingdom: John Wiley & Sons, 2011.

[52] D. Ton, D. C. Duives, O. Cats, S. Hoogendoorn-Lanser, and S. P. Hoogendoorn, "Cycling or walking? Determinants of mode choice in the Netherlands," *Transp. Res. Part A Policy Pract.*, vol. 123, pp. 7–23, 2019.

[53] R. Moeckel, N. Kuehnel, C. Llorca, A. T. Moreno, and H. Rayaprolu, "Agent-Based Simulation to Improve Policy Sensitivity of Trip-Based Models," *J. Adv. Transp.*, vol. 2020, p. 1902162, 2020.

[54] Z. H. Khattak, M. J. Magalotti, J. S. Miller, and M. D. Fontaine, "Using New Mode Choice Model Nesting Structures to Address Emerging Policy Questions: A Case Study of the Pittsburgh Central Business District," *Sustainability*, vol. 9, no. 11. 2017.

[55] "Family Health Outcomes Project: The Planning Guide Appendix III-B." [Online]. Available: https://fhop.ucsf.edu/planning-guide.

[56] L. B. Andersen, P. Schnohr, M. Schroll, and H. O. Hein, "All-cause mortality associated with physical activity during leisure time, work, sports, and cycling to work.," *Arch. Intern. Med.*, vol. 160, no. 11, pp. 1621–1628, Jun. 2000.

[57] J. Woodcock, O. H. Franco, N. Orsini, and I. Roberts, "Non-vigorous physical activity and all-cause mortality: systematic review and meta-analysis of cohort studies," *Int. J. Epidemiol.*, vol. 40, no. 1, pp. 121–138, Feb. 2011.

[58] S. Kahlmeier *et al.*, "Health economic assessment tool (HEAT) for walking and for cycling," Copenhagen, Denmark, 2017.

[59] E. Heinen and K. Chatterjee, "The same mode again? An exploration of mode choice variability in Great Britain using the National Travel Survey," *Transp. Res. Part A Policy Pract.*, vol. 78, pp. 266–282, Aug. 2015.

[60] F. Crawford, "Statistical Release Analyses from the National Travel Survey," 2019.

[61] C. Beltman, "Intrapersonal variability in mode choice behavior: a research based on data from the Dutch mobile mobility panel," University of Twente, 2014.

[62] UK Department for Transport, "Supplementary Guidance: Mixed Logit Models," 2014.

[63] E. Cherchi and C. Cirillo, "A mixed logit mode choice model on panel data: accounting for systematic and random variations on responses and preferences," Feb. 2008.

[64] E. Cherchi and C. Cirillo, "Understanding variability, habit and the effect of long period activity plan in modal choices: a day to day, week to week analysis on panel data," *Transportation (Amst).*, vol. 41, no. 6, pp. 1245–1262, Oct. 2014.

[65] E. Cherchi, C. Cirillo, and J. de D. Ortúzar, "Modelling correlation patterns in mode choice models estimated on multiday travel data," *Transp. Res. Part A Policy Pract.*, vol. 96, pp. 146–153, Feb. 2017.

[66] T. Thomas, L. La Paix Puello, and K. Geurs, "Intrapersonal mode choice variation: Evidence from a four-week smartphone-based travel survey in the Netherlands," *J. Transp. Geogr.*, vol. 76, pp. 287–300, Apr. 2019.

[67] N. Mallig and P. Vortisch, "Modeling travel demand over a period of one week: The mobiTopp model," Jul. 2017.

[68] A. Vij, "Incorporating the Influence of Latent Modal Preferences in Travel Demand Models," University of California, Berkeley, 2013.

[69] “Deutsches Mobilitätspanel (MOP).” [Online]. Available: https://www.bmvi.de/SharedDocs/DE/Artikel/G/deutsches-mobilitaetspanel.html. [Accessed: 23-Aug-2020].

[70] R. Follmer and D. Gruschwitz, “Mobility in Germany – short report,” Bonn: Germany, 2019.

[71] P. Cornick, J. Cant, J. Yarde, C. Byron, and I. Templeton, “National travel survey 2019 technical report,” 2020.

[72] S. Costa, D. Ogilvie, A. Dalton, K. Westgate, S. Brage, and J. Panter, “Quantifying the physical activity energy expenditure of commuters using a combination of global positioning system and combined heart rate and movement sensors,” *Prev. Med. (Baltim).*, vol. 81, pp. 339–344, 2015.

[73] World Health Organization (WHO), “Metrics: Population Attributable Fraction (PAF).” [Online]. Available: https://www.who.int/healthinfo/global_burden_disease/metrics_paf/en/. [Accessed: 14-Oct-2020].

[74] R Core Team, “R: A Language and Environment for Statistical Computing.” Vienna, Austria, 2020.

[75] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.

[76] A. Evans, J. Cummings, M. Slocombe, and F. Corvaglia, “National Travel Survey: England 2017,” 2018.

[77] World Health Organization (WHO), “Physical activity factsheets for the 28 European Union Member States of the WHO European Region. Overview (2018),” World Health Organization, Sep. 2018.

[78] B. Wallmann-Sperlich and I. Froboese, “Physical activity during work, transport and leisure in Germany--prevalence and socio-demographic correlates.,” *PLoS One*, vol. 9, no. 11, 2014.

[79] A. Mok, K.-T. Khaw, R. Luben, N. Wareham, and S. Brage, “Physical activity trajectories and mortality: population based cohort study,” *BMJ*, vol. 365, p. l2323, Jun. 2019.

[80] T. Wei and V. Simko, “R package ‘corrplot’: Visualization of a Correlation Matrix.” 2017.

[81] J. Mullahy, “Specification and testing of some modified count data models,” *J. Econom.*, vol. 33, no. 3, pp. 341–365, 1986.

[82] A. Zeileis, C. Kleiber, and S. Jackman, “Regression Models for Count Data in {R},” *J. Stat. Softw.*, vol. 27, no. 8, 2008.

[83] S. Jackman, “{pscl}: Classes and Methods for {R} Developed in the Political Science Computational Laboratory.” Sydney, New South Wales, Australia, 2020.

[84] S. Hess and D. Palma, “Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application,” *J. Choice Model.*, vol. 32, p. 100170, Sep. 2019.

[85] K. W. Axhausen, M. Löchl, R. Schlich, T. Buhl, and P. Widmer, “Fatigue in long-duration travel diaries,” *Transportation (Amst).*, vol. 34, no. 2, pp. 143–160, 2007.

[86] W. El-Assi, C. Morency, E. J. Miller, and K. N. Habib, “Investigating the capacity of continuous household travel surveys in capturing the temporal rhythms of travel demand,” *Transportation (Amst).*, vol. 47, no. 4, pp. 1787–1808, 2020.

STATEMENT OF INDEPENDENT WORK

I hereby confirm that this thesis was written independently by myself without the use of any sources beyond those cited, and all passages and ideas taken from other sources are cited accordingly.

APPENDIX

This appendix contains the model coefficients for the trip generation models described in chapter 6 and the mode choice models described in chapter 7.

NOTES:

- When an attribute has corresponding values in the second column, the coefficients are usually binary (on/off). However, if the coefficient is multiplied by *n*, then it is multiplied by the value of the attribute (within the range of the corresponding value column).
- When the attribute does not have corresponding values (no second column), the coefficients are always multiplied by the value of the attribute up to the maximum specified.
- In all models, the reference cases are highlighted in light grey.

KEY FOR COEFFICIENTS:

`	*p-value ≤ 0.1*
*	*p-value ≤ 0.05*
**	*p-value ≤ 0.01*
***	*p-value ≤ 0.001*
n	*multiply coefficient by the value of the attribute (within the range given)*

Part A: Trip Generation Model Coefficients

Table A1: Trip generation hurdle model (zero-hurdle part)

attribute	value	HBW	HBE	HBS	HBR	RRT	HBO	NHBW	NHBO
INTERCEPT		1.05***	-3.13***	2.13***	0.25*	-1.03***	1.53***	-3.56***	0.43***
Household size	*1*								
	2	-0.31***		-0.16`	-0.11				-0.20**
	3	-0.46***		-0.20`	-0.32**				-0.59***
	4	-0.47**	0.46***	-0.34*			0.16`		
	5+	-0.58**			-0.60**				-0.92***
Household children	*0*								
	1	-0.36***	-0.43**	0.22*	0.29**		0.65***		0.36***
	2	-0.66***			0.42**				0.45***
	3+	-1.28***	-0.66`		0.67**		1.13***		1.20***
urban	*FALSE*								
	TRUE					-0.23***	-0.18**	0.29***	
Household cars	*0*								
	1	0.40***		-0.25*	0.16`	0.44***		-0.65***	0.24**
	2	0.56***		-0.48***	0.35**			-0.98***	
	3	0.67***		-0.55***	0.49***				
Cars per adult (max. 1)						-0.24*		0.55**	
age	10–18	-3.09***	5.92***	-1.28***	1.08***	-0.24*	-0.52**	-1.58***	
	19–29	-1.29***	2.88***	-0.49***	1.04***		-0.33***	-0.41***	
	30–49								
	50–59	-0.46***	-0.66**	0.19*				-0.48***	-0.43***
	60–69	-2.37***	-0.43`	0.34**	0.51***	0.19*		-0.93***	-0.79***
	70+	-4.11***	-0.20	0.31*	0.53***	0.33***		-1.82***	-1.01***
sex	*male*								
	female	-0.35***		0.16**	0.19**		0.15**		0.27***
driver's license	*No*								
	Yes	0.76***	-0.74***		0.41***		0.55***	0.66***	0.60***
owns bicycle	*No*								
	Yes	0.29***	0.67***		0.41***	0.49***	0.19**	0.21*	0.21***
HBW trips	*0*								
	1–4			-0.19n***		-0.13n***	-0.34n***	0.10n***	0.57***–0.30n***
	5+				-0.15*				
HBE trips	*0*								
	1+			-0.24n***		-0.15n***	-0.31n***	-1.57	0.95***–0.25n***
Primary commute mode	CarD						0.30***	3.88***	
	CarP							3.81***	
	PT						-0.17*	3.72***	
	Cycle			0.47***	0.22*	0.30**		3.79***	
	Walk			0.47***		0.39**		3.49***	
Log usual commute distance (km)					-0.04***		-0.06***		-0.03***

Table A2: Trip generation hurdle model (normal count part)

attribute	value	HBW	HBE	HBS	HBR	RRT	HBO	NHBW	NHBO
INTERCEPT		1.47***	0.74***	0.89***	0.54***	-8.34	0.84***	1.64***	0.99***
House-hold size	*1*								
	2			0.07*			0.08*		
	3			0.12*		-0.21*n****	0.12*		
	4			0.37***			0.19**		
	5			0.51***			0.25**		
House-hold children	*0*								
	1	-0.10***		-0.11*			0.59***	0.24***	
	2			-0.31***	-0.03*n*`		0.72***	0.30***	0.13*
	3	-0.25***		-0.62***			0.87***	0.37**	0.28**
urban	*FALSE*								
	TRUE	-0.04*					-0.10***		
House-hold cars	*0*								
	1								
	2		0.11`	-0.13*n****		0.81***			
	3								
age	10–18	-0.58***	0.70***	-0.55***	0.19***		-0.76***		
	19–29	-0.17***	0.51***	-0.24***	0.25***		-0.34***		0.27***
	30–49								
	50–59		-1.14***	0.10*		0.47***			-0.21***
	60–69	-0.10***	-1.70***	0.17***	0.13***	0.23`		-0.18*	-0.29***
	70+	-0.71***	-2.13***	0.09*	0.25***			-0.52***	-0.58***
sex	*male*								
	female	-0.05***							0.13***
driver's license	*No*								
	Yes		-0.17***	0.08*	0.12**	0.33**	0.29***		0.35***
owns bicycle	*No*								
	Yes				0.20***				
HBW trips	*0*								
	1–4			-0.12*n****	-0.06*n****	-0.13*n****	-0.15*n****	-0.04*n***	-0.14*n****
	5								
HBE trips	*0*								
	1–5			-0.10*n****		-0.41**	-0.17*n****	-0.51***	-0.08*n****
Primary commute mode	CarD			0.11**				-0.49***	
	CarP							-0.36*	
	PT						-0.14**	-0.39***	
	Cycle			0.17***	0.12**			-0.37***	0.11`
	Walk			0.23***	0.17**		0.14*		
Log usual commute distance (km)				-0.02***			-0.02**	-0.02**	
Dispersion parameter log(theta)		17.13	16.52	1.97***	1.25***	-9.40	1.17***	-0.12	0.05

Part B: Mode Choice Model Coefficients

Note: Tables with the attribute "*NESTING PARAMETER*" at the bottom are nested logit models.

Table B1: Usual Commute Mode (McFadden R2 = 0.553)

attribute	value	carD	carP	PT	Cycle	walk
INTERCEPT			1.58**	1.83***	3.07***	5.14***
Household size (max. 5)				-0.34***	-0.15*	-0.31***
Household children (max. 3)				0.44***	0.36***	0.37**
Urban	*FALSE*					
	TRUE			1.03***	0.45***	0.38*
Economic status	*1*					
	2			0.27*	0.28*	
	3				0.5**	
	4		0.31*	0.58**	0.39*	
Cars per adult (max. 1)			-1.43**	-3.91***	-3.7***	-3.27***
PT walk to work/school	≤20min			0.88***	0.51***	0.57***
	>20min					
age	10–18		0.76*	1.42***	0.59*	
	19–29		0.4*	0.72***		
	30–49					
	50–59				0.23*	0.41*
	60–69					
	70+		0.38`			
sex	*male*					
	female		0.28*			-0.26*
driver's license	*No*					
	Yes		-1.80**	-2.36***	-1.78***	-2.21***
owns bicycle	*No*					
	Yes		-0.22*		1.42***	
HBE trips	*0*					
	1–4			0.65***	0.31	0.56*
	5+			1.07***	0.84**	1.46***
Log usual commute distance (km)			-0.13*	0.3***	-1.27***	-2.52***
NESTING PARAMETER		0.46**		0.77***		

Table B2: Mode Restriction Model Multinomial Logit (McFadden R2 = 0.40)

attribute	value	carD carP	carD carP PT Walk	carD carP PT Walk Cycle	PT Walk	PT Walk Cycle
INTERCEPT			1.37***	-2.03***	4.42***	0.04
household children	*0–1*					
	2+		0.42**	0.62***	0.97**	0.76**
Urban	*FALSE*					
	TRUE		0.19*	0.3**	0.66***	0.76***
Economic status	*1*					
	2		0.32**	0.33**		
	3		0.34**	0.49***		
	4		0.37**	0.6***		
Cars	*0*					
	1		-2.47***	-2.27***	-5.33***	-5.68***
	2		-2.67***	-2.70***	-5.92***	-7.06***
	3+		-2.80***	-3.22***		-7.57***
Cars per adult (max. 1)			-0.53*	-0.96***	-1.31*	
PT walk to home	≤10min		0.24*	0.26*	0.53*	0.65**
	>10min					
Age	10–18		1.46***	1.94***	1.90***	1.38**
	19–29		0.41*	0.64***	0.71*	
	30–49					
	50–59			0.25**		
	60–69			0.37***		
	70+				-0.29`	
Sex	*male*					
	female			-0.27***		-0.5***
Driver's license	*No*					
	Yes				-0.36*	
Own bicycle	*No*					
	Yes			1.67***		2.81***
Usual Commute Mode	*carD*					
	carp					
	PT				2.71***	1.09***
	Cycle					0.69*
	Walk			-0.46**	2.91***	
Sqrt HBW trips					-1.47***	-0.41***
Sqrt HBE trips					-1.4***	
Sqrt HBS trips			0.16**	0.45***		0.37**
Sqrt HBR trips			0.46***	0.86***		0.41***
RRT trips (max. 1)				0.21**		0.57***
Sqrt HBO trips				0.23***	-0.58***	-0.51***
Sqrt NHBO trips			-0.24***	-0.33***	-0.29**	-0.49***
Log minimum trip distance (km)			-1.48***	-1.41***	-1.33***	-1.45***
Log maximum trip distance (km)					-0.5***	-0.6***
Coefficient of variation of trip distance			0.45***	0.6***	0.55*	1.06***

Table B3: HBW Nested Logit (McFadden R2 = 0.80)

attribute	value	carD	carP	PT	Cycle	walk
INTERCEPT			-0.33**	-3.98***	1.23***	-0.64`
household size (max. 5)						0.15*
Urban	*FALSE*					
	TRUE			0.81***		
Economic status	*1*					
	2					
	3					0.45**
	4			0.60***		0.55**
Cars	*0*					
	1			-0.44*		-0.59**
	2+			-0.67*		-1.04***
Cars per adult (max. 1)			-0.33***	-1.30***	-1.27***	
PT walk to work/school	≤20min			0.47***	0.31**	0.50***
	>20min					
age	10–18		0.52***			0.91*
	19–29					
	30–49					
	50–59					
	60–69		0.09*			
	70+		0.17*			
sex	*male*					
	female		0.12***	0.26*	-0.29**	
driver's license	*No*					
	Yes		-0.52***	-1.01***	-1.29***	-1.1***
HBW trips	*0-4*					
	5+				-0.17`	
Primary Commute Mode	*carD*					
	carP		1.42***	1.26***	0.97**	0.84*
	PT		0.77***	5.96***	2.93***	2.98***
	Cycle		0.53***	3.19***	4.08***	2.2***
	Walk		0.59***	3.44***	1.91***	4.22***
Log trip distance (km)				0.48***	-0.9***	-1.98***
NESTING PARAMETER		0.25			0.7***	

Table B4: HBE (McFadden R2 = 0.77)

attribute	value	carD	carP	PT	Cycle	walk
INTERCEPT			1.08**	0.88	4.38***	5.55***
household size (max. 5)					-0.34*	
household children	*0–1*					
	2+		0.5**		0.69*	
Urban	*FALSE*					
	TRUE			0.33*	0.5*	-0.68`
Economic status	*1*					
	2					-0.72`
	3					
	4					-2.18*
Cars	*0*					
	1–2			-0.61		-0.91
	3			-0.84`		-2.19*
age	**10–18**		0.83*	0.77*		
	19–29					
	30–49					
	50–59					
	60–69					
	70+					
sex	*male*					
	female		0.28`		0.51*	0.69*
driver's license	*No*					
	Yes		-1.38***	-1.39***	-1.78***	-1.57**
HBE trips	*0-4*					
	5+					-1.04*
Primary Commute Mode	**carD**	3.33***				
	carp		2.13***			
	PT			2.95***		
	Cycle				2.82***	
	Walk					3.61***
Log trip distance (km)				0.3**	-1.23***	-2.87***
NESTING STRUCTURE		[no car nest]		0.86***		

NOTE: due to an insufficient sample size for HBE trips it was only possible to estimate the impact of primary commute mode on the utility of the respective mode.

Table B5: HBS (McFadden R2 = 0.49)

attribute	value	carD	carP	PT	Cycle	walk
INTERCEPT			0.33***	-0.85**	2.73***	2.83***
Household size (max. 5)			-0.07***			-0.15***
Household children	*0–1*					
	2+		0.11**		-0.27*	-0.46***
Urban	*FALSE*					
	TRUE			1.07***	0.13`	0.28***
Economic status	*1*					
	2					
	3		0.04**			-0.31***
	4					
Cars	*0*					
	1			-2.67***	-1.86***	-1.79***
	2+			-2.97***	-2.13***	-2.11***
Cars (max. 3)			0.15***			
Cars per adult (max. 1)			-0.63***	-1.17***		
PT walk to home	≤10min		0.04**	0.78***		0.24***
	>10min					
age	10–18		0.08**	-0.36`	-0.33*	
	19–29					
	30–49					
	50–59		-0.05*		0.43***	
	60–69		0.06**		0.53***	
	70+		0.06*		0.56***	
sex	*male*					
	female		0.33***	0.55***	0.17**	
driver's license	*No*					
	Yes		-0.58***	-0.3*	-0.63***	-0.72***
owns bicycle	*No*					
	Yes			-0.28**	0.17*	
HBW trips	*0*					
	1–4		-0.05**	-0.25`		
	5+			-0.63***		
HBE trips	*0*					
	1+				0.4**	
Usual Commute Mode	carD		-0.05**	-0.41*	-0.36***	-0.38***
	carp		0.21***	0.86**		
	PT		0.15***	1.09***		
	Cycle		0.11***			
	Walk		0.08*	0.64**	-0.41*	0.27`
Log trip distance (km)			0.06***	0.19*	-1.22***	-2.12***
NESTING PARAMETER		0.17		0.73***		

Table B6: HBR (McFadden R2 = 0.38)

attribute	value	carD	carP	PT	Cycle	walk
INTERCEPT			0.53***	0.41	3.52***	3.62***
Household size (max. 5)				-0.11*		-0.15***
Household children	*0–1*					
	2+		0.08*		0.18*	0.31**
Urban	*FALSE*					
	TRUE			1.16***		0.18***
Economic status	*1*					
	2					
	3		0.05**			-0.14**
	4					-0.19**
Cars	*0*					
	1			-1.13***	-1.13***	-0.74***
	2+			-1.51***	-1.38***	-0.83***
Cars (max. 3)			0.1***			
Cars per adult (max. 1)			-0.54***	-1.58***	-0.57***	-0.78***
PT walk to home	**≤10min**			0.43***		
	>10min					
age	**10–18**		0.25**	0.35*		-0.38**
	19–29			0.22*		-0.18*
	30–49					
	50–59				0.17**	
	60–69		0.06**	-0.36***		-0.39***
	70+		0.08*	-0.39***	-0.27***	-0.43***
sex	*male*					
	female		0.4***	0.31***		
driver's license	*No*					
	Yes		-0.93***	-0.9***	-0.85***	-0.82***
owns bicycle	*No*					
	Yes				0.33***	0.21***
HBW trips	*0*					
	1–4		-0.11**			
	5+		-0.1**			
HBE trips	*0*					
	1–4		-0.13**			
	5+		-0.1`		0.27**	
Usual Commute Mode	**carD**			-0.36***	-0.41***	-0.23**
	carP		0.23**		-0.43**	
	PT		0.15**	0.37***	-0.29**	-0.16`
	Cycle		0.21***		-0.12`	
	Walk				-0.53***	0.1
Log trip distance (km)			0.05***	-0.07`	-1.06***	-1.83***
NESTING PARAMETER		0.29***		0.62***		

Table B7: HBO (McFadden R2 = 0.50)

attribute	value	carD	carP	PT	Cycle	walk
INTERCEPT			0.58***	-0.58*	3.13***	2.34***
Household size (max. 5)			-0.08***	-0.1*	-0.14***	-0.13***
Urban	*FALSE*					
	TRUE			1.00***	0.22***	0.31***
Economic status	*1*					
	2				-0.21**	
	3					
	4				-0.35**	
Cars	*0*					
	1			-1.8***	-1.09***	-1.15***
	2+		0.2***	-2.39***	-1.28***	-1.46***
Cars per adult (max. 1)			-0.53***	-1.08***	-0.57**	-0.52**
PT walk to home	≤10min			0.55***		0.28***
	>10min					
age	10–18		0.31***			
	19–29				0.19*	
	30–49					
	50–59					
	60–69			0.18`	0.16`	
	70+					
sex	*male*					
	female		0.19***	0.35***		
driver's license	*No*					
	Yes		-0.52***	-0.85***	-0.69***	-0.85***
HBW trips	*0*					
	1+					0.36***
HBE trips	*0*					
	1+					0.38**
Usual Commute Mode	carD		-0.11***	-0.95***	-0.75***	-0.74***
	carP		0.14***			
	PT			1.13***		-0.39**
	Cycle			-0.39*	0.04	-0.44**
	Walk					0.26`
Log trip distance (km)			0.04***	0.09*	-1.18***	-2.18***
NESTING PARAMETER		0.15		0.71***		

Table B8: RRT (binary logit, McFadden R2 = 0.33)

attribute	value	walk	cycle
INTERCEPT			-3.04***
Household size (max. 5)			-0.23***
Urban	*FALSE*		
	TRUE		-0.21`
age	10–18		1.19***
	19–29		
	30–49		
	50–59		
	60–69		0.26`
	70+		
sex	*male*		
	female		-0.24*
driver's license	*No*		
	Yes		-0.52***
Usual Commute Mode	carD		0.37*
	carP		
	PT		
	Cycle		-0.51**
	Walk		-1.02**
Log trip distance (km)			1.65***

Table B9: NHBW (McFadden R2 = 0.58)

<table>
<tr><th>attribute</th><th>value</th><th>carD</th><th>carP</th><th>PT</th><th>Cycle</th><th>walk</th></tr>
<tr><td colspan="2">INTERCEPT</td><td></td><td>1.34**</td><td>-0.27</td><td>2.78***</td><td>2.59***</td></tr>
<tr><td colspan="2">Household size (max. 5)</td><td></td><td></td><td>0.18**</td><td></td><td></td></tr>
<tr><td rowspan="2">Household children</td><td>0–1</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>2+</td><td></td><td>-0.72***</td><td>-0.53***</td><td>-0.44***</td><td></td></tr>
<tr><td rowspan="2">Urban</td><td>FALSE</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>TRUE</td><td></td><td></td><td>0.72***</td><td></td><td></td></tr>
<tr><td rowspan="4">Economic status</td><td>1</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>2</td><td></td><td>0.58**</td><td rowspan="2">0.94***</td><td rowspan="2">0.34*</td><td>0.47**</td></tr>
<tr><td>3</td><td></td><td>0.99***</td><td>0.57***</td></tr>
<tr><td>4</td><td></td><td>1.24***</td><td>1.54***</td><td>0.4*</td><td>0.94***</td></tr>
<tr><td rowspan="3">Cars</td><td>0</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>1</td><td></td><td></td><td>-0.63*</td><td>-0.94***</td><td rowspan="2">-1.06***</td></tr>
<tr><td>2+</td><td></td><td></td><td>-0.95*</td><td>-1.57***</td></tr>
<tr><td colspan="2">Cars per adult (max. 1)</td><td></td><td>-0.87***</td><td>-1.13***</td><td></td><td></td></tr>
<tr><td rowspan="2">PT walk to work</td><td>≤20min</td><td></td><td></td><td>0.44***</td><td>0.2`</td><td>0.53***</td></tr>
<tr><td>>20min</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="6">age</td><td>10–18</td><td></td><td rowspan="2">0.97***</td><td></td><td></td><td rowspan="2">0.43***</td></tr>
<tr><td>19–29</td><td></td><td></td><td></td></tr>
<tr><td>30–49</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>50–59</td><td></td><td rowspan="3">-0.24*</td><td rowspan="3">0.29**</td><td>0.45***</td><td>0.44***</td></tr>
<tr><td>60–69</td><td></td><td rowspan="2">0.47***</td><td rowspan="2">0.38***</td></tr>
<tr><td>70+</td><td></td></tr>
<tr><td rowspan="2">sex</td><td>male</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>female</td><td></td><td></td><td>0.34***</td><td>0.22**</td><td></td></tr>
<tr><td rowspan="2">driver's license</td><td>No</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Yes</td><td></td><td>-3.17***</td><td>-2.19***</td><td>-2.94***</td><td>-2.23***</td></tr>
<tr><td rowspan="2">owns bicycle</td><td>No</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Yes</td><td></td><td></td><td></td><td>0.45*</td><td></td></tr>
<tr><td rowspan="3">HBW trips</td><td>0</td><td></td><td>-0.37*</td><td>0.72***</td><td>1.29***</td><td>0.50**</td></tr>
<tr><td>1–4</td><td></td><td>0.39***</td><td>0.38***</td><td>0.49***</td><td>0.22**</td></tr>
<tr><td>5+</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="5">Usual Commute Mode</td><td>carD</td><td></td><td>-1.42***</td><td>-2.27***</td><td>-1.22***</td><td>-0.76***</td></tr>
<tr><td>carP</td><td></td><td>1.61***</td><td></td><td></td><td></td></tr>
<tr><td>PT</td><td></td><td>0.69***</td><td>1.95***</td><td>1.33***</td><td>1.89***</td></tr>
<tr><td>Cycle</td><td></td><td>0.36`</td><td>-0.54**</td><td>1.46***</td><td>0.59**</td></tr>
<tr><td>Walk</td><td></td><td></td><td>-0.84***</td><td></td><td>0.74***</td></tr>
<tr><td colspan="2">Log trip distance (km)</td><td></td><td>0.1**</td><td>0.22***</td><td>-0.88***</td><td>-1.98***</td></tr>
<tr><td colspan="2">NESTING PARAMETER</td><td colspan="2">[no car nest]</td><td colspan="3">0.71***</td></tr>
</table>

Table B10: NHBO (McFadden R2 = 0.39)

attribute	value	carD	carP	PT	Cycle	walk
INTERCEPT			0.47*	-0.58`	1.97***	1.47***
Household size (max. 5)			-0.06*	0.17***	-0.12***	
Household children	*0–1*					
	2+			-0.61***		-0.27***
Urban	*FALSE*					
	TRUE			0.51***	0.19***	
Economic status	*1*					
	2		0.04`			
	3		0.06*			-0.2***
	4			0.18`	0.21*	0.31***
Cars	*0*					
	1		0.08	-1.47***	-0.99***	-0.99***
	2+		0.23*	-1.6***		
Cars per adult (max. 1)			-0.49*	-1.08***	-0.89***	-0.64***
PT walk to home	≤10min		-0.03`	0.55***	0.32***	0.32***
	>10min					
age	10–18		0.24*	0.59***		0.57***
	19–29		0.06`	0.48***		0.42***
	30–49					
	50–59		-0.06*			
	60–69					
	70+					0.2***
sex	*male*					
	female		0.23*	0.23***		0.1*
driver's license	*No*					
	Yes		-0.62*	-0.67**	-0.68**	-0.72***
HBW trips	*0*					
	1+		0.08*			0.19***
HBE trips	*0*					
	1+			0.50***	0.62***	0.35***
Usual Commute Mode	carD		-0.12*	-0.74***	-0.64***	-0.36***
	carP		0.16*	0.26`		
	PT		0.09*	0.72***		0.38***
	Cycle			-0.43***	0.15*	
	Walk			-0.34*		0.19`
Log trip distance (km)			0.04*		-0.86***	-1.29***
NESTING PARAMETER		0.2*		0.59***		