ASSOCIATE PROFESSORSHIP FOR TRAVEL BEHAVIOR

TUM SCHOOL OF ENGINEERING AND DESIGN

TECHNISCHE UNIVERSITÄT MÜNCHEN

Modelling Access and Egress Connections to Transit Services in England

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

Advisors: Dr. Qin Zhang
Author: Ghassan Ammar
ghassan.ammar@tum.de
Submission Date: March 27, 2023

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.

Munich, March 27, 2023

Ghassan Ammar

Ghassan Ammar

Acknowledgments

I would like to start by expressing my immense gratitude to my supervisor, Dr. Qin Zhang, for her steadfast support and guidance. Her availability to answer my inquiries at all times is deeply appreciated. Furthermore, I would like to extend my sincerest thanks to Corin Staves, who ably assisted me with the UK national travel survey data and imparted valuable insights on a number of topics.

I am also deeply beholden to the esteemed members of the Travel Behavior Chair, particularly to Prof. Dr.-ing Rolf Moeckel, who offered me the opportunity to be a part of the Chair, and to Dr. Ana Moreno and Dr. Carlos Llorca, who provided me with invaluable assistance and direction during my tenure.

My sincerest gratitude goes to my mother, Amal, who has been a constant source of love, support, and encouragement throughout my academic journey. I am deeply thankful to my grandparents, Geli and Bernd, for their wisdom, guidance, confidence in my abilities, and for being a shining example of resilience. I am grateful for all the sacrifices, love, and support you have provided me over the years. Without you, this journey would not have been possible, and for that, I am eternally thankful.

Contents

1. Introduction

1.1. Background

Intermodal travel, which refers to the integration of various modes of transportation within a single journey [1], has become a crucial aspect of transportation planning in recent years. This is driven by several factors, such as the overburdening of transport infrastructure in urban areas and the steady growth of the population [2]. The expansion of motility options in terms of both quantity and diversity has made multimodal trips more attractive [3], however, it has also made the movement patterns more intricate [4].

The primary goal of evaluating public transportation accessibility is to improve the connection between people and locations, thereby reducing congestion on the roads. In other words, the use of public transportation can help to mitigate the negative impacts of private vehicle usage on the environment and public health [5]. For this reason, it is important to consider the accessibility of public transportation, the interconnectivity of different modes of public transportation, and the overall transportation system to create a user-friendly public transportation system [6].

The increasing focus on sustainable development has highlighted the importance of accessibility as a key metric to evaluate transport investments and urban policies [7], therefore, urban planners are now placing more emphasis on public transportation over private transportation [8]. Access and egress have a wide-ranging impact on various aspects of society, including daily life, public health, social inclusion, sustainability, economic efficiency, spatial efficiency, temporal efficiency, and the environment [9].

Despite the wealth of research on mode choice, there is limited focus on characterizing multimodal trips, particularly with regard to the role of access and egress [10, 11]. These modes, often seen as the weakest link in travel journeys, are frequently ignored in mode choice analysis. [12, 13]. The availability of public transportation, largely determined by access and egress, has a significant impact on its utilization [14]. To utilize the system, users must be able to reach the transfer station and continue to their final destination. To effectively use the system, users must be able to reach the transfer station and continue to their final destination. A well-functioning line-haul system becomes useless if access and egress to the transfer station and final destination are problematic, impacting the perceived accessibility of public transportation [15].

1.2. Modelling Access and Egress

Transport models, which have their roots in the 1950s, were traditionally focused on predicting road congestion for private car travel. However, as the emphasis on sustainable transportation has grown, there has been an increased interest in modelling active modes of travel such as walking and cycling [16]. Improvements in transport modelling aim to better understand the full effects of transportation decisions, however, traditional transportation models typically only evaluate mobility and fail to account for accessibility. To address this, there has been a growing interest in multimodal approaches that take into account the quality of non-motorized and public transportation options, as well as integrated transportation and land-use models that focus on accessibility.

Encroachments in computational power have led to the development of agent-based models, which are better equipped to understand the full implications of transportation decisions. These models are not restricted to practical data and can simulate scenarios that may not be feasible to conduct in reality. They are stochastic models that are constructed from the bottom up, where individual agents are assigned specific characteristics. The agents are then programmed to interact and behave in certain ways with other agents and their surroundings. Agent-based models are more versatile and allow for the integration of statistical models with a broader array of predictors, making it possible to study complex systems that display the interdependence of individuals [17].

Access and egress modelling has the potential to provide valuable insights that can assist policymakers and transport agencies to optimize their design and operation, ultimately reducing wait times and refining traffic flow. By analysing data on mode choices, trip characteristics, and other factors that influence access and egress, organizations could gain an enhanced understanding of the flow of people and vehicles, identifying areas where improvements could be made. For instance, by ascertaining bottlenecks or other inefficiencies in access and egress systems, facilities can make targeted improvements to reduce congestion and improve safety.

Another advantage of access and egress modelling is the potential for increased accessibility for individuals with disabilities and other mobility challenges. Through analysing the impact of design and operational factors on accessibility, organizations can identify areas where improvements could be made to better accommodate all users.

1.3. Research Question

This research aims to improve the design and accessibility of transit systems by gaining insights into the key factors that influence travellers' behaviour and preferences when accessing and egressing public transport in England. Specifically, the study aims to answer the research question: **What are the key factors influencing the choice of travel mode for commuters and non-commuters accessing and egressing public transport in England?**

The study seeks to optimize ridership and enhance accessibility through the development of a predictive model that incorporates socio-demographic, spatial, and journey characteristics indicators. By providing a deeper understanding of the decision-making process involved in choosing travel modes, the study can serve as a reliable reference for improving transit systems. The research's potential to contribute to the development of more efficient and sustainable transportation systems has significant implications for urban planning and environmental policy.

1.4. Overview of this Thesis

Chapter 2 of this thesis provides a comprehensive review of the literature on modelling access and egress modes, with a focus on both separate and joint modelling approaches. Additionally, it examines influential factors that have been found to impact travel mode choice in general, as well as specifically on access/egress mode choice.

In Chapter 3, the chosen data sources are described, and the steps taken to prepare the dataset for modelling are discussed, including a discussion of key limitations and important changes that were made. This chapter also outlines the methodology used to model access and egress modes, and proposes a modelling framework that aims to improve the effectiveness of these models.

Chapter 4 presents the results of the modelling analysis, including an explanation of the findings and projections of potential real-world scenarios to illustrate the implications of the results. In addition, this chapter offers a broader discussion of the findings and their implications for future research. Chapter 5 provides a comprehensive conclusion that summarizes the key findings of the study and offers a discussion of its limitations and broader implications for future research on this topic.

Finally, The regression summary tables for the models are provided in the Appendix section to facilitate a more comprehensive understanding of the statistical analyses performed.

2. Literature Review

A comprehensive literature review was performed to examine the methods used in prior studies for modelling access and egress. The review emphasized the importance of access and egress as often-neglected components in travel sequences and revealed their significant impact on various aspects beyond transportation, especially when choosing the access/egress mode. The review examined the key parameters that influence access and egress in theoretical and modelling studies and provided a detailed analysis of how other researchers have modelled access and egress. This last section was divided into three parts: studies focusing solely on access, studies focusing solely on egress, and studies covering both access and egress.

2.1. Importance of Access and Egress in Intermodal Transport

Tønnesen et al. [18] highlighted the significance of multi-modal analysis, noting that while it has gained increasing importance in transport planning, the areas of access, egress, and transfer when using public transportation are still under-researched. The study emphasized the benefits of enhancing public transportation by improving access, egress, and transfers. They proposed that incorporating active modes of transportation such as biking or walking for access and egress in conjunction with public transportation would not only decrease the use of private cars and subsequently lower carbon emissions but also have a positive impact on public health compared to solely relying on private cars for door-to-door travel.

The authors in [18] also pointed out that there exists a competitive relationship between different modes of transport. This means that travellers tend to choose the mode of transportation they perceive to be the most appealing. Therefore, they suggested that land use and transport infrastructure should be designed in a way that makes sustainable modes such as walking and cycling more appealing and gives them a competitive edge over private cars.

Pucher et al. [19] covered in their study the impact of daily walking and cycling captured in transport trips on public health, and they concluded that these forms of active travel contribute to daily physical activity, aerobic fitness, and cardiovascular health while helping to protect against obesity, diabetes and various other diseases. Van Soest et al. [20] confirmed also in their literature review the positive influence of access/egress trips on foot on health. The authors in [19] also added that the mounting evidence on the health benefits of walking and cycling has led many governments' medical journals to advocate more walking and cycling to improve individual health and reduce air pollution, carbon emissions, congestion, noise, traffic dangers, and other harmful impacts of car use. In short, there is consensus on the need to increase daily walking and cycling levels to promote public health.

The availability and ease of access and egress play a crucial role in determining an individual's choice of using public transportation. In their research, Zhen et al. [21] evaluated the effect of access and egress on overall satisfaction with rail transit and found that the travel time to access the rail transit negatively affects satisfaction, while familiarity with the access route increases satisfaction. The results for gender and age are consistent with those for satisfaction with rail egress. The authors found that egress travel time also has a negative impact on satisfaction, but familiarity with the egress route and the use of electronic devices to search for information on reaching the final destination are positively associated with satisfaction.

The availability and ease of access and egress play a crucial role in determining an individual's choice of using public transportation. In their research, Zhen et al. [21] evaluated the effect of access and egress on overall satisfaction with rail transit and found that the travel time to access the rail transit negatively affects satisfaction, while familiarity with the access route increases satisfaction; the authors found that egress travel time also has a negative impact on satisfaction. Gutiérrez et al. [22], Cervero [23], and Keijer and Rietveld [24] found that longer access and egress times contribute to a more negative experience while travelling, leading to a decrease in the likelihood of travellers choosing to use public transportation; in other words, when access and egress times become excessively lengthy, it acts as a significant deterrent to the use of public transportation.

Brons et al. [25] found that the satisfaction of passengers with their rail journey is largely dependent on the quality of the access facilities provided to them. Hence, enhancing access to railway stations is expected to drive up rail usage. Their research shows that access facilities hold even greater significance for infrequent rail travellers, suggesting that improving access to the rail network can not only attract new passengers but also encourage greater use of rail services among these individuals.

2.2. Modelling Studies

2.2.1. Significant Parameters

De Witte et al. [26] discussed in their paper a conceptual framework (Figure 1) for identifying and structuring modal choice determinants, the framework uses a multi-disciplinary approach including economics, transport geography, and social psychology. The framework has two circles: the first circle distinguishes 3 types of determinants (socio-demographic, journey characteristic, and space-related) that shape options for making a modal choice, while the second circle represents the influence of socio-psychological factors (habits, perceptions, experiences) that determine how these options are acted upon.

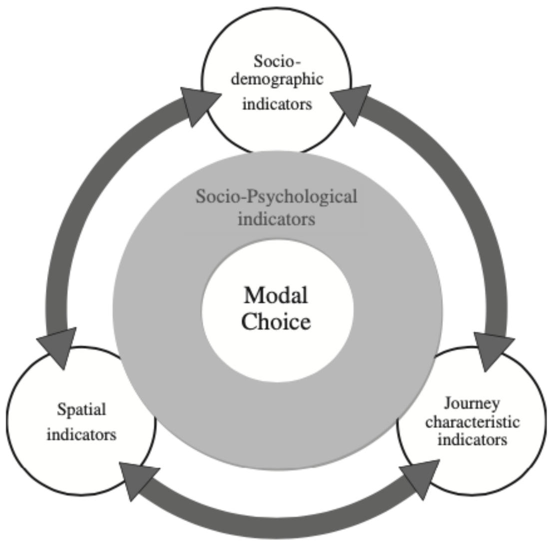

Figure 1. Framework for structuring modal choice determinants (source: De Witte et al. [26])

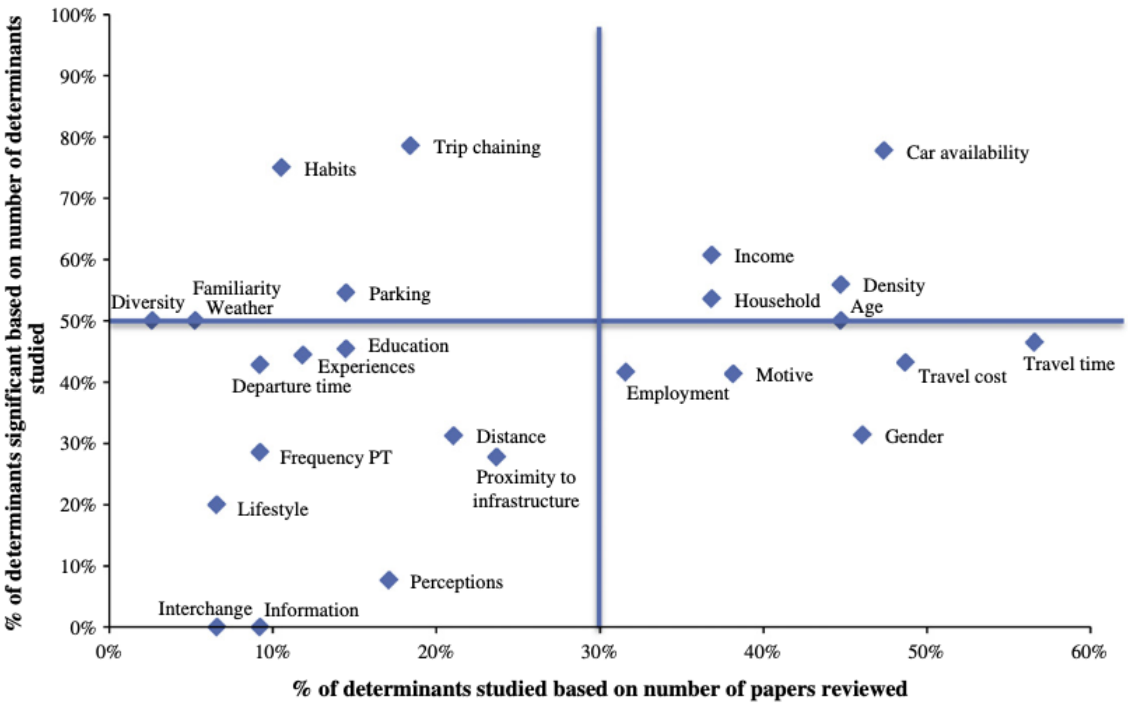

Figure 2. Classification of modal choice determinants based on the review analysis (source: [26])

De Witte et al. study aims to evaluate the importance of determinants in modal choice decisions by identifying the determinants considered in different papers and verifying their significance. The results are shown in a classification (Figure 2), where the horizontal axis indicates the ratio of papers that include a determinant to the total number of papers reviewed, and the vertical axis indicates the ratio of times a determinant is found significant to the number of times it is studied. Based on the representation, determinants can be roughly classified into four groups: rarely studied and infrequently found significant, rarely studied and frequently found significant, often studied and rarely found significant, and frequently studied and usually found significant. Objective and straightforward determinants, like socio-demographic characteristics and car availability, are more frequently studied compared to subjective and complex ones, like familiarity and lifestyle. Some of the subjective and complex determinants, however, tend to have a significant influence on the modal choice decision.

Studies [13, 25, 27–29] have emphasized the role of personal and household characteristics in access/egress mode choice, particularly the sociodemographic factors. Kim et al. [13] found significant impacts of gender, age, driver's license, employment status, and household income on transit access choice; females were more likely to use the bus as an access mode, with gender also affecting private vehicle availability, crime rates, and time of day. Loutzenheiser [30] noted that men were more likely to walk than women, while Tran et al. [29] did not find gender to significantly impact slow mode usage. Age and gender were found to have a significant effect on rail egress satisfaction, with older passengers and women having higher satisfaction according to Zhen et al. [21].

The effect of car availability was studied by Givoni and Rietveld [10] and Kim et al. [13], with the latter finding an increased likelihood of the drive&park alternative. However, Givoni and Rietveld did not find a strong effect of car availability on access mode choice. Kim et al. [13] found younger travellers to be more likely to be picked up/dropped off, while older travellers favoured slow modes. A driver's license reduced the likelihood of using the bus or being picked up/dropped off, and car ownership positively impacted car access and negatively impacted public transport and walking [28, 30]. Lower household income correlated with increased bus use and walking [13, 28, 30], though Tran et al. [29] couldn't confirm this effect. Other factors that affected access choice, according to Tran et al., were the number of children and workers, while Loutzenheiser [30] cited education level as a factor in walking to access stations.

Land use has a significant impact on access mode choice, as noted in studies by Kim et al. [13], Krygsman et al. [27], and Wen et al. [28]. Residents of urban areas are more likely to walk than those living downtown. Additionally, numerous studies emphasize the significance of station proximity (to and from) as the main factor in choosing an access/egress mode [13, 23, 27–30]. As expected, the likelihood of non-motorized modes decreases with increasing distance to the transit station. Jiang et al. [31] specifically investigated the effect of the built environment on walking access mode choice and concluded that people are willing to walk for long distances to the station in a busy and interesting environment.

Cervero and Duncan [32], Zhao and Li [33], and Krygsman et al. [27] have also investigated the influence of factors related to the built environment on people's choices of transport and access/egress modes. The results of these studies indicate that micro-level elements such as the density of the population and places of employment, the mixture of land uses, and the transportation infrastructure near a station, as well as the starting point of the journey, all play a role in shaping mode choices. The location of the station and the presence of parking facilities in its vicinity are also crucial factors to consider.

Polydoropoulou and Ben-Akiva [34] analysed transport-system-specific elements in determining both access and main public transit mode and found factors like the number of transfers, in-vehicle travel time, cost of parking, transit fare, walk access time and delay probability to be significant. Wen et al. [28] also found similar factors impacting access mode choice. Arentze et al. [35] discussed the empirical estimation of individuals' preferences in their selection of multimodal travel options stating that the time spent on access and egress, as well as wait and walk time, is considered more important than in-vehicle time, but not to the extent that is often thought. Trip purpose had also a significant impact, as found Wen et al., with costly modes like taxis being preferred for work-related trips over recreational trips. Weather conditions were highlighted as a contextual factor by Krygsman et al. [27].

2.2.2. Choice Models in Transport

Choice models are applied in transportation and other fields to capture the choice of a single option from a set of exclusive alternatives. The logit family (multinomial, nested, cross-nested, and mixed) has a convenient closed-form formulation, making estimation easier [36]. Even though the multinomial logit (MNL) model (McFadden, 1973) is the most popular choice model and is known for its straightforward mathematical structure and ease of estimation [37]; it is often criticized for the Independence of Irrelevant Alternatives (IIA) property of its error terms, which restricts choice in certain situations [36, 38].

The IIA property assumes that if someone can't drive, their probability of choosing other modes of travel will increase evenly; however, factors like the distance to their destination could actually increase the likelihood of taking public transit over walking or biking [39]. Using the MNL model in such a situation leads to incorrect results. To address the IIA issue, the nested logit models were created, removing the IIA property from alternatives to nests. The nested logit model provides a better alternative as it allows for specified mode pairs to show increased sensitivity to changes in the service quality [36, 37] since alternatives with correlated errors are grouped together in a nest.

2.2.3. Modelling Studies

Access Studies

Bergman et al. [40] used discrete choice modelling methods to examine the behaviour of individuals when choosing between alternative access modes. The study developed two methods to calculate the utility of variables related to the mode (e.g., travel time, cost, and frequency), the decision maker (e.g., income, car ownership, and age), and the environment (e.g., population density). The first method is a simple multinomial logit (MNL) model while the second is a nested logit (NL) model which allows for different degrees of similarity between alternative access modes. The study utilized survey data gathered onboard during the first year of operation of WES (Westside Express) in the Portland Oregon metropolitan area, as collected by Tri-Met, the regional transit agency.

Bergman et al. research [40] concluded that the NL model with the best log-likelihood score of -679.47 outperformed the multinomial logit (MNL) model (p <.0001) in predicting the mode of transportation choice. The rho-squared relative to 0 was 0.324 and the adjusted rho-squared relative to constants was 0.146, considered acceptable considering data limitations. The authors evaluated multiple nesting structures before choosing the structure in Figure 3 based on fit, parameter estimates and theoretical validity. In this structure, car, bus, and LRT were grouped together as they share more unexplained variance, while drop-off differs from the motorized modes. Similarly, biking and walking were placed in the same nest as closer substitutes.

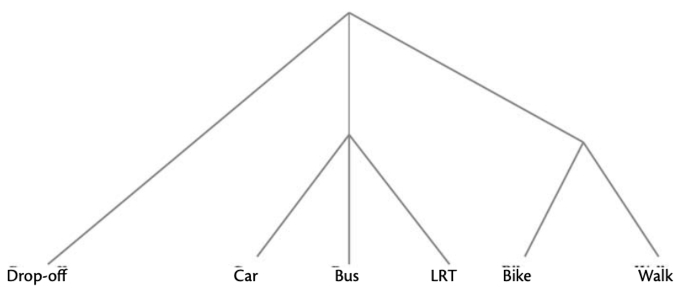

Figure 3. Preferred nesting structure, (source: Bergman et al. [40])

Debrezion et al. [41] modelled access mode and railway station choice jointly, they analysed departure railway station and access mode choices made by Dutch travellers, utilizing the comprehensive Railway Service Quality Index (RSQI) as a station feature. A nested logit model was applied to explain the choice behaviour where alternatives with the same access mode has been grouped together. The model was estimated using data from 1440 postcode areas and various access and railway station attributes.

The study [41] starts by assuming that the passenger has already made the decision to travel by train, which is influenced by various factors including railway station accessibility. A more thorough analysis would involve considering all modes of transportation; however, the data only reflects train trips. The passenger is then presented with two interconnected choices - the method of accessing the station and the departure station - which are both made simultaneously. The initial analysis used a multinomial logit model with 12 options; however, this model did not take into account the correlation between choices. To address this, the authors utilized a Nested Logit model and analysed two structures - one where access mode is at the top level and departure station at the lower level, and the other in reverse order. The inclusive value parameters from the estimation helped determine the most appropriate nesting structure. The results indicated that the first nesting structure was the best fit for the analysis (Figure 4).

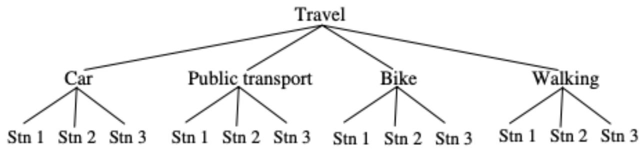

Figure 4. Access mode departure station choice decision tree, (source: Debrezion et al. [41])

Findings

Bergman et al. [40] did not find a significant relationship between walk access propensity and distance to the nearest transit station. The other findings of their research indicate that factors such as the number of car parking spaces, the availability of connecting bus routes, and population density at a station are major predictors of the mode of access to WES. The study also discovered a correlation between certain riders' attitudes and their mode of access to WES, with a preference for comfort leading to a higher likelihood of accessing WES by LRT or car, and a pro-sustainability mindset resulting in a preference for bike access. The estimation results showed that low-income riders only access WES through cost-free bus transport. Furthermore, the study provides clear evidence that former car users prefer driving and parking at rail stations, and employer-sponsored pass holders are more inclined to access WES by car.

Debrezion et al. [41] findings indicate that all variables have a significant impact on the selection of both access mode and departure station. Distance has a negative impact on the choice of departure station, with a steeper impact observed for non-motorized modes of transportation such as walking and biking, suggesting they are preferred for shorter distances. Car ownership has a positive but insignificant effect on the choice of car access and a negative effect on the use of public transportation. The availability of parking spots and bicycle stands has a positive impact on the choice of departure stations accessed by car and bicycle respectively. Public transport frequency has a positive impact, while public transport travel time has a negative impact on the choice of departure stations accessed by public transport. The RSQI (Rail Service Quality Index) of a station has a significant and positive effect on the choice of departure stations for all modes. The choice probability elasticity with respect to the RSQI increases as one moves from the first station with the highest share to the third station with the lowest share for all access modes, with the elasticities being generally higher for motorized modes compared to non-motorized modes.

Egress studies

Molin and Timmermans [11] conducted a study on 996 train travellers to show that context variables, including travel purpose, time of day, weather, travel party, amount of luggage, distance, and route knowledge, have a significant impact on their mode of egress choice. The authors conducted two experiments to gather these results. The first experiment varied the choice alternatives, while the second experiment varied the context variables to arrive at a set of context descriptions. The choice sets from the first experiment were then nested under the context descriptions from the second experiment, resulting in a set of context-choice sets equal to the number of choice sets multiplied by the number of context descriptions. Since the analysis aimed to assess the effect of context variables on the selection of egress mode, the authors suggested that a simple MNL was appropriate for this purpose.

In their study, Meng et al. [42] aimed to gather information on last-mile home-bound trip makers for each mode using quota sampling at five major rail transit stations in different parts of Singapore. The goal was to randomly obtain at least 50 respondents for each mode (cyclists, pedestrians, and others) in each station of the five major stations in Singapore. To analyse mode choice decisions, the study proposed the use of a multinomial logit (MNL) modelling approach. The mode choice was defined as a dependent variable and modelled using a generalized logit approach with SAS software.

The study [42] defined three dependent variables: the probability of choosing to walk, cycle, or take a public feeder bus as a last-mile home-bound trip maker. The final model was selected through automatic variable selection in SAS software and manual adding of variables. The variables that were considered were actual distance travelled, presence of integrated transport hub, availability of personal household vehicle, age, number of bikes along intermediate links surrounding the transit station, and gender. The significance of each variable was checked by adding them one by one, starting with the variable with the greatest x2 and significant p-values. Interactions among variables were also considered.

Findings

Molin and Timmermans study [11] aimed to investigate the factors that influence egress mode choice, with seven alternatives considered: public transport (PT), taxi, train taxi, public PT bike, bike in train, bike at station, and Greenwheels (rental car based on shared car principles). The study found that context variables, such as trip purpose, knowledge of the route, weather, distance to final destination, amount of luggage, time of day and travel party, have an impact on egress mode choice and were divided into two levels. The study also found evidence of socio-demographic differences in the influence of context variables, with gender differences found to be particularly significant. These results have implications for policy assessment and suggest that including gender and other socio-demographic variables in models of egress mode choice decision-making can reduce heterogeneity in decision-making.

Meng et al. [42] found that the most significant factors influencing mode choice for last-mile trip stages are the actual distance between the transit station and destination and the number of bicycles along intermediate links surrounding the transit station. Secondary factors include socio-demographic variables such as age, gender, and household income. Tertiary factors include the number of feeder bus services and vehicle availability. In general, travellers tend to prefer walking for shorter distances, cycling for longer distances, and using public buses for the furthest distances. The number of bicycles along intermediate links is positively associated with walking and cycling modes. The results also showed that older travellers are more likely to choose cycling and that males are more likely to walk and cycle compared to females. Households with an income of less than $2,000 tend to cycle rather than take a bus, and the non-availability of a private vehicle increases the likelihood of walking and cycling.

In Between Access and Egress

Kurth et al. [43] explained that in 1989, Metra, a railway operator, conducted a survey on mode of access for rail lines users in the Chicago area. The survey was self-administered on trains and included questions on the mode of exit and final destination. This produced a wealth of data with 10,741 individual observations that were used to estimate central area travel models. The average walking time for egress trips on foot was 12.4 minutes, which equates to approximately 0.6 miles. This is significantly longer than the commonly used 0.33-mile walk distance used in many regional modelling processes.

In their study, Schakenbos et al. [44] investigated the inconvenience of transferring between local public transport feeder modes (bus, tram, and metro) and the train as the primary mode in a public transport trip. They estimated the value of access and egress time and observed a clear distinction between bus trips for access and egress; they found that VOT is higher for access/egress compared to the in-vehicle part. Puello et al. [45] reported that egress trips are perceived worse than access trips indicating that VOT is higher for egress. Hensher et al. [46] found that people have higher VOT for access than for the main part of the journey but this does not hold for the egress part, with the exception of car trips.

Rahman et al. [47] found that the choice of access and egress modes is contingent upon the choice of the main mode of transportation, with different scenarios emerging for access and egress. Their study's results suggest that the main mode choice significantly influences the egress mode choice, with rail travellers more likely to choose motorized transportation at the end of their journey. However, the model estimation did not find the choice of access mode to be a significant explanatory factor.

Access and Egress Studies

Azimi et al. [48] together with the LYNX transit agency in the Orlando metro area carried out an onboard survey from January to April of 2017. The study collected 13,181 responses from within the LYNX service area. The survey gathered comprehensive information about all aspects of transit trips, including the access and egress links, as well as the demographic and household information of the users. Most trip origins were located near streets and roads and were geocoded using street GIS data.

The survey data in [48] included 13 access/egress modes, which were grouped into 6 major modes as shown in Figure 5. Transportation Network Company (TNC) and taxi trips were combined into a single category, as TNC trips made up a very small share, and the services are similar. The data was complemented by the land use attributes obtained from the Smart Location Database (SLD), which were incorporated into the database. This information includes population and employment information, density and diversity measures, land use design variables, transit-related attributes, and accessibility measures at the census block group level.

Mode Group	Access/Egress Mode	Access Trip		Egress Trip	
		Frequency	**Percentage**	**Frequency**	**Percentage**
Walk	Walk	83,076	90.5%	84,811	92.4%
TNC or Taxi	Taxi Uber, Lyft, etc.	295	0.3%	481	0.5%
Micromobility	Personal Bike Bike share Skateboard Scooter	2060	2.2%	2377	2.6%
Drive Alone	Drove alone and parked Car share (e.g., Zip Car, etc.)	1700	1.9%	1186	1.3%
K&R, Carpool, Shuttle	Was dropped off or picked up by someone Drove or rode with others and parked Shuttle	4535	4.9%	2809	3.1%
Wheelchair	Wheelchair	122	0.1%	123	0.1%
Total		91,787	100.0%	91,787	100.0%

Figure 5. Mode category, (source: Azimi et al. [48])

The authors in [48] developed two MNL models to investigate the mode choice for access and egress links. In their two models, walking was designated as the base category and the utilities for the other five alternatives were calculated relative to walking. To ensure that the explanatory variables in the MNL models were uncorrelated, the authors conducted correlation tests between the independent variables. Of the correlated variables, only one was included in the analysis (the specific variable was not mentioned).

The fully finalized models for both access and egress in Azimi et al. study [48] showed a lower AIC and BIC value than the initial models, indicating that they performed acceptably. The Akaike's Information Criterion (AIC) is a method that uses in-sample fit to predict future values [49], while the Schwarz Criterion (SC) or Bayesian information criterion (BIC) measures the trade-off between model fit and complexity [50]. Both full models showed improved goodness-of-fit compared to the initial models, and this was supported at a 5 per cent significance level by the log-likelihood ratio test (LRT) (Figure 6).

Items	Access Model		Egress Model	
Criterion	Initial Model	Full Model	Initial Model	Full Model
AIC	11512.294	9914.894	9707.009	6528.688
SC	11549.645	11857.156	9744.36	7537.171
−2 LOGL	11502.294	9394.894	9697.009	6258.688
LL	2107.4004		3438.3205	
Max-Rescaled R-Square	0.255		0.442	

Figure 6. Model performance results, source: Azimi et al. [48])

In their research, Puello et al. [45] have combined revealed and stated preference surveys to study access and egress mode choice to railway stations in the Netherlands. 1524 respondents completed a revealed preference survey and were then subjected to an adaptive stated choice experiment as part of an online survey. The authors focus on walking and cycling as the two most important public transport feeder modes, which account for approximately 60 per cent of the total access and egress travel.

The stated preference data was modelled using mixed logit ML models, which the authors in [45] considered the most suitable for this type of experiment. The models included cost and time factors as well as variables that describe the quality of stations and station environment. In the stated choice access experiment, four attributes were considered: time, cost, and the status of pedestrian and cycling facilities. Five alternatives were included: car, BTM (Bus-Tram-Metro), walk, bicycle, and no choice. The authors divided the 'no choice' option into two: 'I would not travel by train' or 'I would find another way to go to the station'. In the egress experiment, the alternatives were BTM, OV-fiets, bicycle (own), walk, and no choice. The authors tested several specifications before arriving at the final model specifications, and insignificant variables were removed from the model specification.

Unlike Azimi et al. and Puello et al., Creemers et al. [51] used data from the Flemish National Household Travel Survey to jointly predict access/egress and main public transit mode and did not carry their own survey. The data was selected to focus only on public transit journeys and was filtered using three rules:

1. walking trips under 10 minutes were neglected.
2. A public transit journey was considered unimodal if there was no access or egress.
3. For journeys that were not considered unimodal, they were defined as multimodal.

The main access/egress mode for each journey was determined based on environmental impact, with car being prioritized. If car was not used, BTM was considered the access/egress mode, but only in cases where the main mode was train. If neither car nor BTM was used, slow modes were considered the access/egress mode. (Figure 7)

The authors in [51] used a multinomial logit MNL model and the journeys with BTM as the main mode was chosen as the reference category. The results of the model can be interpreted as the impact on the log-odds ratio of each alternative mode compared to the reference Uni_BTM. The model shows a satisfactory fit with a Nagelkerke R2 index of 0.64 and a McFadden R2 of 0.30. This is supported by the results of the likelihood ratio lack-of-fit test, where the null hypothesis indicating a lack of fit was rejected with a P-value of less than 0.001.

Uni/multi	**Main transport mode**	**Access/egress mode**	**Label**	**Observed Frequency**	**Percentage (%)**
Unimodal	BTM	/	Uni_BTM	1067	48.46%
	Train	/	Uni_Train	51	2.32%
Multimodal	BTM	Car	Multi_BTM_Car	98	4.45%
		Slow	Multi_BTM_Slow	336	15.26%
	Train	Car	Multi_Train_Car	276	12.53%
		BTM	Multi_Train_BTM	162	7.36%
		Slow	Multi_Train_Slow	212	9.63%

Figure 7. Descriptive results of public transport journeys, source: Creemers et al. [48])

Mo et al. [52] conducted a comprehensive study on the impact of the built environment (BE) on people's first- and last-mile travel mode choices (i.e. access/egress); moving from the traditional modelling of mode choice into incorporating certain variables besides mode of transport. Singapore was selected as the case study area as its residents heavily rely on public transportation for daily travel, with 70 per cent of commuters using public transit during morning peak hours, according to the 2012 Household Interview Travel Survey. The study aimed to understand the role BE plays in daily travel behaviours, especially in the context of Singapore.

The BE data used in the study [52] was obtained from the Singapore Land Authority digitized cadastral dataset and synthetic population. The authors divided Singapore into 1,169 traffic analysis zones (MTZs) for the calculation of BE variables. The study assumed that the first- and last-mile travel mode choices are influenced by three categories of factors: sociodemographic characteristics of respondents, BE at the origin, destination, and non-MRT station areas, and trip-specific variables. A mixed logit (ML) model framework was used to capture the heterogeneity of the impact of BE.

The authors used two binary ML models (with and without BE variables) to estimate the impact of BE on the first- and last-mile travel mode choices. The model showed a high goodness-of-fit value with an improved adjusted r2 from 0.733 to 0.832 after incorporating BE variables, indicating the importance of BE in addition to trip-specific and sociodemographic variables. The study also adopted a multinomial MNL model to model the impact of BE in areas with LRT and found a high goodness-of-fit value of 0.885 with BE variables and a substantial decrease in goodness-of-fit after discarding BE variables.

Rahman et al. [47] aimed to investigate the modes of transportation used by travellers during their first and last miles of public transit in Dhaka, Bangladesh and the factors affecting their access-egress mode choice behaviour. To obtain answers to these questions, the authors conducted an onboard survey and found that the respondents utilized a range of modes such as walking, rickshaws, paratransit options (human hauler, shared auto-rickshaws), and buses. The majority of the travellers relied on non-motorized transportation (NMTs) including walking and rickshaws.

The study in [47] applied discrete choice models to analyse the access-egress mode choice behaviour of public transit users. The authors compared Multinomial Logit (MNL) and Nested Logit (NL) models before choosing the most suitable model that represents the observed behaviour. Various possible nesting structures were evaluated to establish the relationships between alternative modes. The NL Model was also compared to MNL and produced higher log-likelihood values for both access and egress compared to MNL. To assess the statistical significance, the authors performed log-likelihood ratio tests under the assumption that there was no significant improvement in the model. For egress, although the LR was comparatively lower, the NL still displayed better model fit statistics than MNL ($P < 0.10$). As a result, for both access and egress, the NL model (Figure 8&9) is more appropriate for predicting the mode choice of this dataset than the MNL specification.

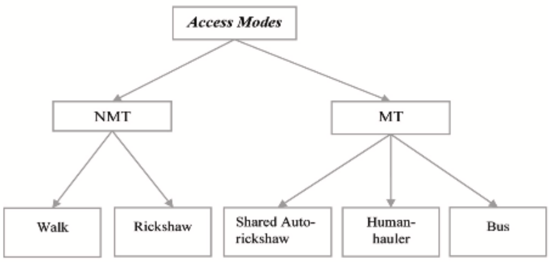

Figure 8. Access model nest structure [47]

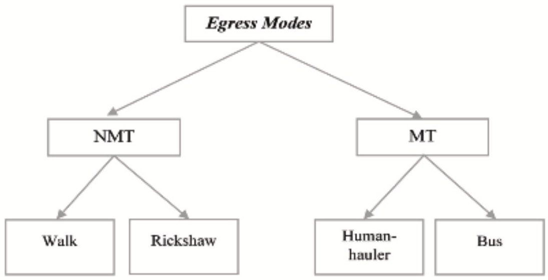

Figure 9. Egress model nest structure [47]

Halldórsdóttir et al. [53] departs from previous research by examining access and egress mode choices not based on the direction of travel, but instead on the location of travel. The focus is on whether the train stations are at the home or activity ends. This approach provides a more consistent understanding of travel mode preferences, for example between a bike ride to the station from home and the return bike ride from the same location. Their study proposes a unique perspective, suggesting that travellers don't have differing preferences for accessing or leaving train stations, but rather have differing preferences at the home-end and activity-end of their train trips. The hypothesis is that these differences arise due to the varying availability of travel modes between the home and activity location, and the greater familiarity with the road network, parking options, and station characteristics at the home end compared to the activity end.

The authors in [53] specifically focus on analysing access and egress to train stations in the Copenhagen Region, utilizing data from travel diaries collected in the Danish National Travel Survey. The analysis employs a joint model comprised of two mixed logit specifications that consider the choice of travel mode at the home and activity end for access and egress to train stations. The study includes 2,921 home-end and 3,658 activity-end trips and considers five alternative modes of travel: walking, biking, driving as the driver, riding as a passenger, and taking the bus. The two mixed logit models were first estimated independently, and the best specification was determined by evaluating the significance of both fixed and random parameter estimates. The best specification for both models showed significant error components indicating the interdependence between active modes of travel (walking, biking) and motorized modes of travel (driving, riding as a passenger, bus). Then, the two models were estimated jointly, and the hypothesis of equal parameters between the home-end and activity-end models was tested.

Findings

Azimi et al. [48] presented the following findings of their access model. As the access length increases and the number of transfers required grows, people tend to opt for non-walking modes such as micro-mobility, taxis or ride-hailing services, carpooling, or driving alone; indicating a significant effect of distance to transit station and number of transfers variables. It was also found that young adults tend to walk to transit stations and mid-income users prefer micro-mobility while high-income users prefer driving or carpooling, indicating the significant effect of age, and household income. Employment and household entropy increase the use of micro-mobility, and high residential density discourages the use of all modes except walking.

Regarding the egress model in Azimi et al. research, the study concluded that trip purpose plays an important role in choosing transport mode, where users going to medical visits or hospitals tend to drive from transit station to their destination, while university students were less likely to carpool. Also, walking was more popular for egress trips to shopping places. The number of transfers had a negative impact on the use of motorized modes for egress purposes compared to walking. Carpool and drive alone were less likely to be used in the mid-day period, indicating that they are typically used for work trips during AM or PM peaks. Gender shows a significant impact as male users were more likely to walk on their egress trips compared to their female counterparts. Regarding household income, both mid and high-income household users were more likely to choose micro-mobility or TNC or taxi. Lastly, Private car-related modes were less likely to be used at destinations with a high number of households with zero auto ownership.

Puello et al. [45] concluded that pedestrians are more sensitive to travel time compared to cyclists during the access journey, while cyclists are more sensitive to travel time compared to pedestrians during the egress journey. Both walking time and travel time by bicycle act as key predictors in the modal choice to access the station, indicating that both pedestrians and cyclists are more sensitive to variations in travel time than car drivers or bus users. A delay in travel time is more relevant for pedestrians compared to cyclists during the access stage. Bicycle costs also have a significant influence on the "no choice" selection. The effect of bicycle parking costs and access time by bicycle was tested as a ratio parameter (time/cost).

The results of the study [45] indicated that trip purpose plays a significant role in the mode choice where bicycle costs have a greater influence on modal choice during work journeys compared to non-work journeys. The perceived penalty of the cost of parking the bicycle at the station slightly increases for those who travel for work purposes, which is contrary to the authors' expectations. Delays are less relevant for non-work journeys compared to work journeys, and people are less flexible and willing to spend more time travelling by car or bike when the journey is for work purposes. The contribution of bicycle cost is greater in the model of work journeys compared to the model of non-work journeys.

All in all, Puello et al. study [45] found that non-work journeys have smaller standard deviations in terms of the panel effect of car and BTM (Bus-Tram-Metro) users, indicating fewer socioeconomic factors influence travel decisions; while there is a greater influence of random factors in the model for non-work journeys. The results also showed differences in socioeconomic and level-of-service attributes in the egress part, with delays being irrelevant in the case of bicycle use and consistent results for pedestrians in both the access and egress parts. The influence of socioeconomic factors and individual preferences is stronger for work journeys compared to non-work journeys.

Creemers et al. [51] concluded that the majority of socio-demographical factors have a significant impact on public transport mode. However, household size, the number of cars, and the presence of children in a household did not have a noticeable effect. Despite prior research indicating that the number of cars can significantly affect access/egress mode [28, 30, 48], this was not observed in this study, which may be due to the model controlling for driver's licenses and car availability. Additionally, the impact of having a partner could not be supported by literature as no studies discussed in the literature review examined this factor.

Creemers et al. also added that access/egress and main public transit mode choice are significantly influenced by factors such as journey purpose, urbanization degree, car availability, and the total distance of the journey. The most significant factor is distance, with the highest chi2-value and the same degrees of freedom. Car availability is the second strongest determinant, indicated by the second highest chi2-value and the same degrees of freedom for all attributes. Surprisingly, the origin of the journey starting at the home location did not have a significant impact, even though different transport modes were expected to be available at home compared to activity locations. This lack of significance may be due to the effect of car availability.

The binary ML model in Mo et al. [52] study suggests that sociodemographic variables have little effect on the first- and last-mile mode choice, meaning that self-selection may not have a significant impact. Commuting-related trips (e.g. work and education) are more likely to be taken by bus. The model highlights that travel time is a key factor in mode choice, and individuals are more sensitive to time spent on the bus than on foot. Distance to MRT station has a significant impact on modal choice, with a longer distance resulting in a higher likelihood of choosing the bus. A higher density of building floor space in non-MRT areas has a negative effect on bus mode choice, suggesting that more socioeconomic activities encourage people to walk. Only the walking-based EAI (Ease of Access Index) to MRT station and floor space density vary across individuals, indicating varying preferences for walking and socioeconomic activities among the sample.

The results of the multinomial ML model in [52] show that bus travel time, distance to MRT station, and EAI (Ease of Access Index) to bus stop have a robust impact on modal choice behaviour, as reflected in the binary ML model. People have no inherent preference for bus or LRT modes when LRT is available. The model suggests that people tend to use LRT when they are further from MRT station, as the increased distance may encourage more people to use LRT over bus. The model also shows that high levels of land-use mix (entropy) will encourage more people to walk.

The results of the model estimation in Rahman et al. [47] show that factors such as time and cost have a negative impact as expected. Waiting and transfer times for vehicles lead to increased discomfort compared to travel time of the same duration. Additionally, the effect of in-vehicle time and out-of-vehicle time (OVT) differs among modes of transportation. Motorized mode users are less affected by OVT compared to rickshaw users, while travellers using human haulers and buses are willing to endure longer waiting times. On the other hand, users of non-motorized transportation, such as walking and rickshaws, are more sensitive to travel time than motorized transport users. Walking, as it involves physical exertion, is particularly burdensome, whereas using a rickshaw comes with added cost. Therefore, a one-unit increase in travel time would have a greater negative impact on users of non-motorized transportation than on motorized transportation users.

Rahman et al. study also indicates that users of motorized transportation are more sensitive to travel costs than those using non-motorized transportation. Socioeconomic factors such as income, gender, age, education, and occupational status also play a role in choosing access and egress modes. Higher-income travellers are more likely to use rickshaws, while female travellers are more likely to choose rickshaws and less likely to choose human haulers and buses. Age, education, and occupational status, however, were found to have little effect.

The findings of Halldórsdóttir et al. [53] revealed that improving neighbourhood features, cycling infrastructure, and bus services could have a significant impact on the use of active travel modes for accessing train stations. The results also showed that the design of train stations themselves can play a role in encouraging active travel. Specifically, increasing the number of covered parking spaces for bicycles was found to have a positive impact, as it would nearly triple the probability of using a bicycle to access the train station. On the other hand, adding parking spaces for bicycles at the home-end of the trip was not found to have a positive effect. Furthermore, policies restricting the use of bicycles on trains, either due to time restrictions or additional costs, were found to have a negative impact on the choice of using a bicycle at the activity-end.

The study [53] also shed light on the relationship between traveller characteristics, trip purposes, and access and egress mode choices. It was found that existing policies, such as reduced fares for students and the elderly, have contributed to a preference for bus use among these groups. However, the study also identified specific preferences among other traveller groups, such as those travelling for shopping and leisure.

2.3. Research Contribution

European cities have an advantage over their UK counterparts in terms of ease and speed of public transportation to city centres. 67 per cent of residents in major European cities can reach the city centre within 30 minutes by public transport, whereas only 40 per cent of people in large British cities have the same access [54]. The poor state of urban transportation and high car dependence in UK cities, which are among the most car-reliant in Europe [54], has become a growing concern among policymakers, researchers, and citizens alike.

The literature reviewed in this thesis highlights the significant impact that improving the access and egress experience can have on the preference for public transportation over private transportation. While the existing literature on access and egress presents common themes and challenges faced by cities around the world, including Europe, to the author's knowledge, no literature has delved into the subject of access and exit in England with the same level of depth as other -European- areas.

Despite the extensive discussion of context's impact on decision-making in the literature (e.g. Ariely and Levav [55]), the effect of context on mode choice has received limited attention and is frequently neglected in mode choice applications. Most existing models either ignore context or only take into account a single context, such as the commute to work, resulting in limited practicality [11]. To gain a comprehensive understanding of mode choice in multi-modal chains, it is crucial to investigate the influence of contextual factors such as weather and travel companions on mode choice decisions [11]. Furthermore, the impact of the built environment on mode choice for access and egress has been shown to be influential, yet it is not thoroughly explored in many literary works.

The purpose of this study is to fill the knowledge gap mentioned in the preceding paragraph to the greatest extent possible given the available data. Furthermore, this thesis will redirect its focus from the conventional analysis of trip purpose to a physical trip analysis, which could offer valuable insights into the field of Physical Activity (PA) modelling. The novelty of this study lies in its attempt to integrate contextual factors and infrastructure satisfaction factors within the framework of an activity purpose focus, as it investigates access and egress mode choices on a significant scale.

This thesis draws upon the existing literature to address the identified research gaps in the following manner:Firstly, it will compare and evaluate different datasets to determine the most appropriate for modelling access and egress in England. Secondly, it will undertake a rigorous data processing phase to select and validate relevant variables that align with the research objectives. Lastly, it will employ a combination of descriptive and statistical techniques to analyze the impact of various factors on access and egress choices.

Significant variables for access mode	Confirming Studies
Distance to transit station	Mo et al., Azimi et al., Debrezion et al., Krygsman et al., Kim et al., Loutzenheiser., Wen et al., Tran et al., Cevero.
Land use and infrastructure	Mo et al., Puello et al., Halldórsdóttir et al., Krygsman et al., Loutzenheiser., Jiang et al., Cervero and Duncan, Zhao and Li, Brons et al.
Travel time	Mo et al., Puello et al., Bergman et al., Debrezion et al., Rahman et al.
Household income	Azimi et al., Bergman et al., Rahman et al., Kim et al., Loutzenheiser., Wen et al.
Age	Azimi et al., Kim et al., Loutzenheiser., Tran et al.
Gender	Bergman et al., Rahman et al., Kim et al., Loutzenheiser.
Trip Purpose	Creemers et al. Mo et al., Halldórsdóttir et al. , Wen et al.
Density of the neighbourhood	Azimi et al. Creemers et al., Mo et al., Bergman et al.
Possession of driving license	Azimi et al. Creemers et al., Kim et al.
Number of car/bike parking spaces available at transit station	Bergman et al., Debrezion et al., Halldórsdóttir et al., Brons et al.
Cost	Puello et al., Bergman et al., Rahman et al.
Transport mode	Wen et al., Polydoropoulou and Ben Akiva.
Number of vehicles in a a household	Azimi et al., Debrezion et al.
Car Availability	Creemers et al., Kim et al., Givoni and Rietveld.
Time of the day	Azimi et al., Kim et al.
Ease of access to transit station	Mo et al.
Number of transfers	Azimi et al.
Employment rate	Azimi et al.
Delay	Puello et al.
Possession of transit pass	Bergman et al.
Public transport frequency	Debrezion et al.
Service quality of transit station	Debrezion et al.
Weather	Krygsman et al.

Table 1. Significant variables for the access mode

Significant variables for egress mode	Confirming Studies
Trip Purpose	Azimi et al., Puello et al., Creemers et al., Mo et al. Molin and Timmermans, Halldórsdóttir et al.
Number of transfers	Azimi et al.
Time of the day/trip	Azimi et al., Molin and Timmermans, Kim et al.
Gender	Azimi et al., Molin and Timmermans, Meng et al., Rahman et al., Kim et al., Loutzenheiser., Zhen et al.
Household income (income)	Azimi et al., Meng et al., Rahman et al., Kim et al., Loutzenheiser., Wen et al.
Cost	Puello et al., Rahman et al.
Delay	Puello et al.
Travel time	Puello et al., Mo et al., Rahman et al.
Socioeconomic	Puello et al.
Possession of driver's license	Creemers et al., Kim et al.
Car availability	Creemers et al., Meng et al.
Urbanization degree	Creemers et al.
Travel distance	Creemers et al.
Distance from transit station	Mo et al., Molin and Timmermans, Meng et al., Krygsman et al., Kim et al., Loutzenheiser., Cevero.
Density of the neighbourhood	Mo et al.
Ease of Access to transit station	Mo et al.
Land use mix and infrastructure	Mo et al., Halldórsdóttir et al., Krygsman et al., Loutzenheiser., Cervero and Duncan, Zhao and Li
Weather	Molin and Timmermans, Krygsman et al.
Availability of bikes at transit station	Meng et al., Halldórsdóttir et al.
Age	Meng et al., Kim et al., Loutzenheiser., Tran et al., Zhen et al.

Table 2. Significant variables for the egress mode

3. Data Sources & Methodology

3.1. Data Sources

In today's world, data is vital to decision-making processes. The ability to analyse data to extract meaningful insights has become progressively important crosswise all industries. Therefore, choosing the right dataset for a specific analysis is crucial to ensuring precise and reliable results. For this thesis, two datasets were considered. The first dataset is The Greater Manchester Travel Diary Surveys (TRADS), and the second dataset is the National Travel Survey (NTS).

Greater Manchester Travel Diary Surveys

The Greater Manchester Travel Diary Surveys (TRADS) is a program designed to gather transport and travel information from a sample of 2,000 households per year. The purpose of the survey is to collect data pertaining to all trips made by residents aged four years and above during a 24-hour period. The sampling methodology aims to ensure that each district within Greater Manchester is represented proportionally based on the demographics of the population. The program spans the duration of one year, with surveys conducted every day. The survey program yields data on approximately 7,000 trips made by 4,500 residents of 2,000 Greater Manchester households annually. The information collected encompasses trip origins and destinations, travel durations, modes of transportation employed, and the purpose of the journey [56].

National Travel Survey (England)

The National Travel Survey (NTS) is a household survey that serves the purpose of tracking long-term trends in personal travel and guiding policy development. NatCen Social Research conducts the NTS on behalf of the UK Department for Transport. It is the leading data source for personal travel patterns by individuals residing in England within Great Britain. The survey is a comprehensive travel study that dates back to 1965 and has been a continuous survey since 1988, with annual implementation. On average, the survey interviews approximately 8,500 households each year. Although some households may participate in multiple years, their unique identifiers are not retained in the data, so each year's records must be treated as independent. The NTS boasts a substantial sample size, broad coverage, and meticulous measurements and is documented over a seven-day travel diary [57, 58].

Chosen Dataset

The use of TRADS in documenting zonal trips has advantages. Zonal systems divide geographic areas into zones for easier analysis and are especially useful in transportation planning and modelling as they allow for the comparison of trip patterns across different regions. However, NTS provides more detailed documentation of staged trips, including origin, destination, and intermediate stops, which is beneficial for simulating access and egress trips. Moreover, the NTS covers the entire country of England, as opposed to TRADS which only covers Manchester, allowing for greater generalizability of findings to the wider population. Conversely, even though TRADS documents origin, destination, and trip purpose, it lacks staged documented trips, making it less useful for analysing trip patterns, especially in urban areas where trips tend to be more complex. Therefore, for the present thesis, the National Travel Survey (NTS) was preferred over The Greater Manchester Travel Diary Surveys (TRADS). This study solely utilised NTS data collected starting 2014 as a substantial modification in the data collection technique took place in that year.

3.2. Selected Attributes and Filtering

The selection of attributes for the statistical models in the current thesis was based on a careful review of the literature and the data available in the dataset. Through a comprehensive literature review, a number of attributes were identified that could potentially impact access and egress trips, such as trip purpose, mode of transportation, distance, time of day, and demographic characteristics of the traveller, among others. The selection of these attributes was guided by their relevance to the research question and their importance in previous studies on travel behaviour.

Furthermore, the examination of the dataset was crucial in selecting the attributes for the statistical models. Each attribute was evaluated for data quality, and any attributes with significant data quality issues were excluded. In addition, the consistency and reliability of the selected attributes were considered to ensure that they were accurately and consistently measured throughout the dataset.

3.2.1. Included Attributes

The relevant variables for each category are given in Tables (3- 4- 5).

Variable Description	Variable Name in The Dataset
Unique ID	PSUID
Survey year	SurveyYear
PSU Country	PSUCountry_B01ID
Household region	PSUStatsReg_B01ID
Household ID	HouseholdID
Number of people in a household	HHoldNumPeople
Number of bicycles in a household	NumMCycle
Number of cars in a household	NumCar
Household income	HHIncome2002_B02ID
Household urban	Settlement2011EW_B03ID
Household area classification	HHoldOAClass2011_B03ID
Satisfaction with walking infrastructure	Pavement_B01ID
Satisfaction with cycle infrastructure	CycLane_B01ID
Time to the nearest bus station	WalkBus_B01ID
Time to the nearest rail station	WalkRailAlt_B01ID
Walk time to the nearest food store	WalkTimeGroc_B01ID
Walk time to the nearest shopping centre	WalkTimeShopC_B01ID
Walk time to the nearest doctor	WalkTimeGP_B01ID
Walk time to the nearest post office	WalkTimePO_B01ID
Walk time to the nearest chemist	WalkTimeChem_B01ID
Walk time to the nearest hospital	WalkTimeHosp_B01ID
Satisfaction with bus service (in the region)	SatServ_B01ID
Reliability of rail service (in the region)	RelMetro_B01ID

Table 3. Primary Sampling Unit & Household Variables

Variable Description	Variable Name in The Dataset
Person ID	PersNo
Age	Age_B01ID
Gender	Sex_B01ID
Ownership of driving licence	DrivLic_B02ID
Ownership of season ticket	TicketHolding_B01ID
Ownership of bicycle	OwnCycle_B01ID
Work status	EcoStat_B03ID

Table 4. Persons Variables

Variable Description	Variable Name in The Dataset
Travel day of the trip	TravDay
Sequence of journeys	JourSeq
Trip origin	TripPurpFrom_B01ID
Trip destination	TripPurpTo_B01ID
Travel mode	MainMode_B03ID
Travel distance	TripDisIncSW
Trip time	TripTotalTime
Mode used for the stage	StageMode_B03ID
Stage weight	SSXSC
Stage distance	SD
Stage time	StageTime

Table 5. Trips & Stages Variables

3.2.2. Harmonising and Preparing the Data

Designated Primary Travel Mode	NTS Data
Car Driver (carD)	Household car–driver Non-household car–driver Household motorcycle driver Non-household motorcycle driver Private(hire)bus Household van/lorry–driver Non-household van/lorry passenger Household van/lorry passenger Non-household van/lorry passenger Other private transport
Car passenger (carP)	Household car–passenger Non-household car passenger Household motorcycle passenger Non-household motorcycle passenger
Cycle	Bicycle
Walk	Walk, less than 1 mile Walk, 1mile or more

Table 6. Standardizing the Modes of Travel

Designated Activity Purpose	Designated Purpose	NTS Data
Mandatory Activities	Work Education	Work In course of work Education
Escort Activities	Accompanying	Escort home Escort work Escort in course of work Escort education Escort shopping/personal business Other escort
Discretionary Activities	Shopping Recreation Other Non-home based work	Food Shoppig Non-food shopping Eat / drink with friends Visit friends Other social Entertain / public activity Sport: participate Day trip Just walk Personal business medical Personal business eat/drink Personal business other Holiday: base Other non-escort

Table 7. Standardizing the Trip Purpose

After analysing the data, it was determined that the variables outlined in light grey in Table 3 had not been accurately captured within the National Travel Survey (NTS) data. These variables were initially included with the objective of developing a new household area classification system to measure serviceability in the area and assess its impact on mode choice. However, due to the inaccurate capture of these variables, they were deliberately excluded from the subsequent analysis to ensure the reliability and validity of the results.

Variables that are shaded in a darker grey in the same table were documented differently starting in 2014 and were not present in the available dataset. Consequently, in an effort to capture the fullest possible range of data, the dataset was expanded to include records between 2002 and 2013. However, the resulting number of records containing information about walk time to the nearest bus and/or train station was deemed insufficient for reliable modelling, and these data were subsequently excluded from the estimation. As a result, only data from 2014 onwards were used to construct the model.

The purpose of trips was redefined from a destination-oriented approach to an activity-based approach. Mandatory activity trips, which include work and education trips, were classified as one type of activity purpose. Escort activity trips, which include accompanying trips, were also defined as a separate type of activity purpose. Discretionary activity trips, encompassing all other trip purposes such as shopping and recreation trips, were likewise identified as a distinct category of activity purpose. Furthermore, to restrict the scope of this thesis to short trips, a classification system was implemented that categorized trips into two categories: long trips and short trips. Long trips were defined as mandatory trips that exceeded 100 kilometres and were subsequently excluded from the analysis. Similarly, escort and discretionary trips were considered long trips if they exceeded a distance of 40 kilometres and thus were also omitted from the analysis.

The dataset was analysed to identify stage clusters that corresponded to variables from trips, persons, and households to define access and egress for each trip, ensuring non-duplication. In this context, access was defined as the first stage in trips that had a minimum of two stages recorded, subject to two conditions:

1. The access mode must be one of car driver, car passenger, bus, cycle, or walk.
2. The main travel mode must be a PT mode, specifically, either train or bus.

Egress, on the other hand, was defined as the final stage in trips with a minimum of three stages recorded, using the same aforementioned criteria.
The NTS data inadequately represents other modes that could be used for access and egress beyond those that were explicitly defined. As such, the selection of access and egress modes was limited to the aforementioned ones to ensure the reliability and validity of the analysis. Similarly, the selection of PT modes was motivated by their wider geographical spread and higher representation in the dataset compared to other PT modes. Trips that utilized private modes of transportation as the main mode were excluded from consideration in both the access and egress stages.

Following the harmonization and filtering of the NTS dataset, Table 8 reports the resulting descriptive statistics.

Descriptive Statistics	Dataset (all trips)	Models' Dataset (filtered -not weighted- trips)
No. of Households	152793	15806
No. of Persons	363639	24778
No. of Trips	4937043	187966
No. of stages	5168225	187966
No. of PT trips with 2 stages at least	-	21484
No. of PT trips with 3 stages at least	-	9737

Table 8. Summary Statistics: Multimodal and Non-Multimodal Trips

3.2.3. Understanding Access-Egress Data

The modal split and travel behaviour of individuals are critical components of transportation planning and policy. This part examines the modal split and travel characteristics of access and egress modes, including walking, cycling, and motorized transport modes. The aim is to provide insights into the patterns of travel behaviour, including the time and distance associated with each mode, as well as the percentage of individuals using different modes.

Figures 10 and 11 display the modal split for access and egress modes, respectively. The primary mode for both access and egress is walking, followed by the bus. Private transportation modes, such as driving a car or being a passenger, are more prevalent as access modes for both mandatory and discretionary activity trips compared to their egress counterparts. Additionally, the dataset reveals that bicycle usage percentages for both access and egress are very low and do not exceed 3-4 per cent. Notably, the walk mode dominates as the primary mode for both access and egress trips.

Figures 12 and 13 depict the travel time and distance for each mode of access and egress. In the first boxplot, it is evident that walking has the lowest median access travel time (7 minutes) when compared to other modes, all of which have a median value of 10 minutes. For egress travel time, both walking and cycling modes exhibit shorter travel times compared to motorized modes. Notably, car trips (either as a driver or a passenger) show higher values for access and egress travel times than other modes combined. The second box plot shows that access distances tend to have relatively higher median values than their egress counterparts, with car trips as a driver having the longest distance travelled among all modes. The walking mode has considerably shorter access and egress distances in comparison to biking trips, while the egress trip as a passenger in a car has the shortest distance travelled among all motorized modes.

Figures 14 and 15 pertain to the access travel time and distance to the primary public transport modes, i.e., bus and train, while figures 16 and 17 relate to the same parameters for egress. According to the boxplots, walking represents the mode of transport with the shortest travel time and distance to public transportation (PT) stations, with a median travel time of 10 minutes or less and a median distance of 400m to 800m. Cycling to train and bus stations also represents a relatively fast option, with median travel times of 10 minutes and a median distance of 2.4 km to 3.7 km. However, driving as either a passenger or driver entails longer travel times, with median travel times ranging from 8 to 15 minutes, which can be attributed to the longer travel distance of between 3 km to 8 km. Taking the bus to the train is the slowest option, with a median travel time of 10 minutes, while taking the bus to the bus has a median travel time of 21 minutes and a median distance of 3 km to 4.8 km. It is noteworthy that the relatively longer travel time on the bus compared to the shorter distance could be explained by the operational characteristics of buses, which have to cover numerous stops in adjacent areas, as opposed to taking the shortest available route.

The egress travel time box plot highlights that the median egress travel time for most modes of transport is between 5 to 10 minutes, with the exception of egress travel time by car as a driver from a bus, which has a median of 2 minutes. Conversely, the egress distance box plot illustrates that the median egress distance for most modes of transport is between 400 to 800 meters, with the exception of egress distance by cycle from train, which has a median of 2 km and the longest range for egress distance by car as a driver from train (4.8 km). Furthermore, the data demonstrates that cycle trips as an egress mode are non-existent when the main travel mode is bus. This observation can be attributed to various factors, such as the lack of bike racks or dedicated parking spaces close to bus stops or the limited space on buses to carry bikes while travelling.

Figures 18 and 19 illustrate cumulative trips for access and egress, categorized by distance. The data presented in both figures are consistent and in line with the expected outcomes. The analysis indicates that approximately 90 per cent of walking trips cover a distance of roughly 1.5 km for both access and egress. On the other hand, cycling and bus trips primarily cover a distance of 7.5 km for access and a slightly shorter distance for egress. Conversely, car trips, whether as a driver or a passenger, exhibit a higher distance value, ranging from around 10 to 12 km for access and between 6 and 7.5 km for egress. It is noteworthy that access trips cover longer distances for all modes, except for walking.

The analysis of the modal split for access and egress modes, as well as the associated travel time and distance, has provided valuable insights into the travel behaviour of individuals in England. The data shows that walking is the primary mode for both access and egress, followed by the bus. Private transportation modes are more prevalent as access modes compared to egress modes. The data also shows that bicycle usage percentages for both access and egress are very low.

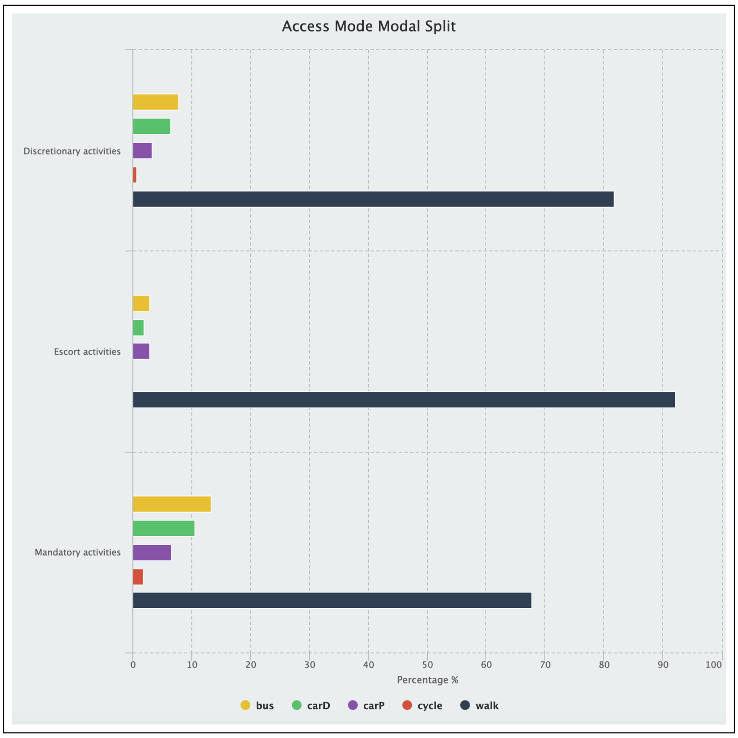

Figure 10. Access Mode Modal Split

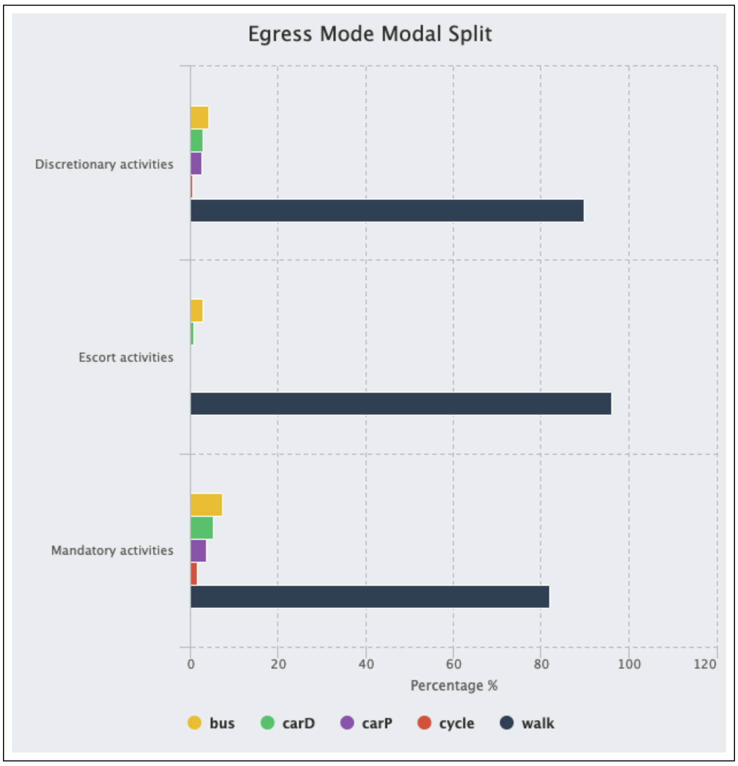

Figure 11. Egress Mode Modal Split

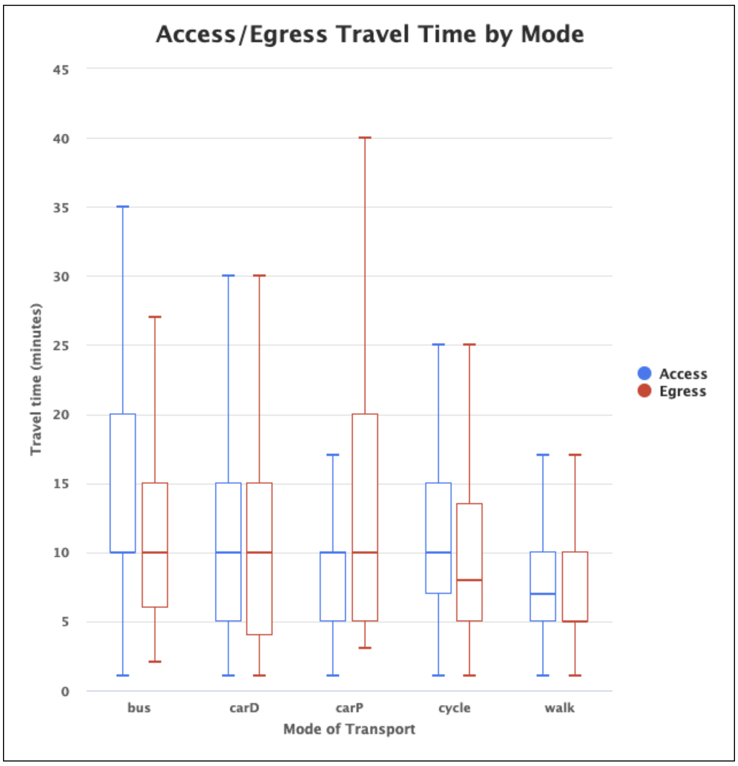

Figure 12. Access/ Egress Travel Time by Mode

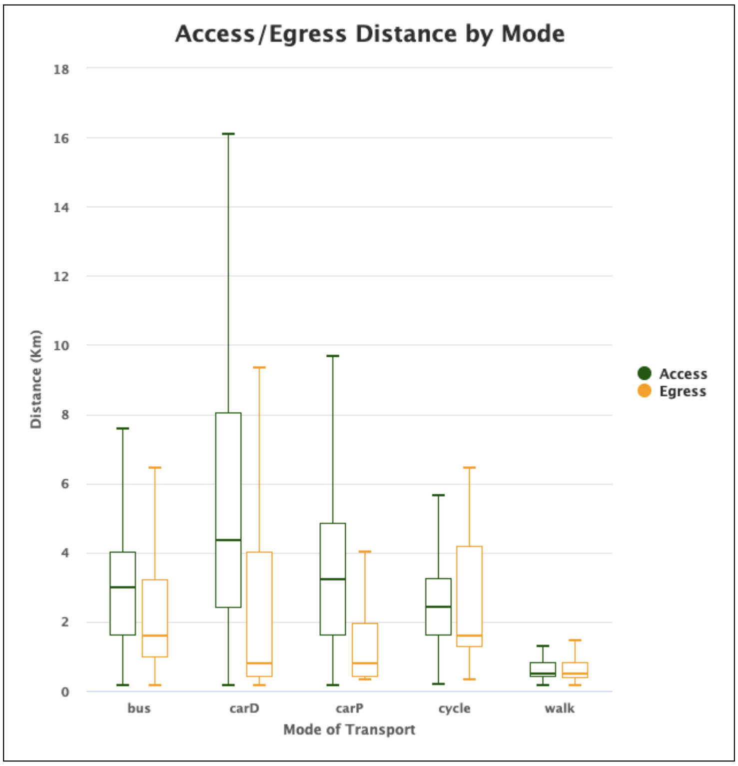

Figure 13. Access/ Egress Distance by Mode

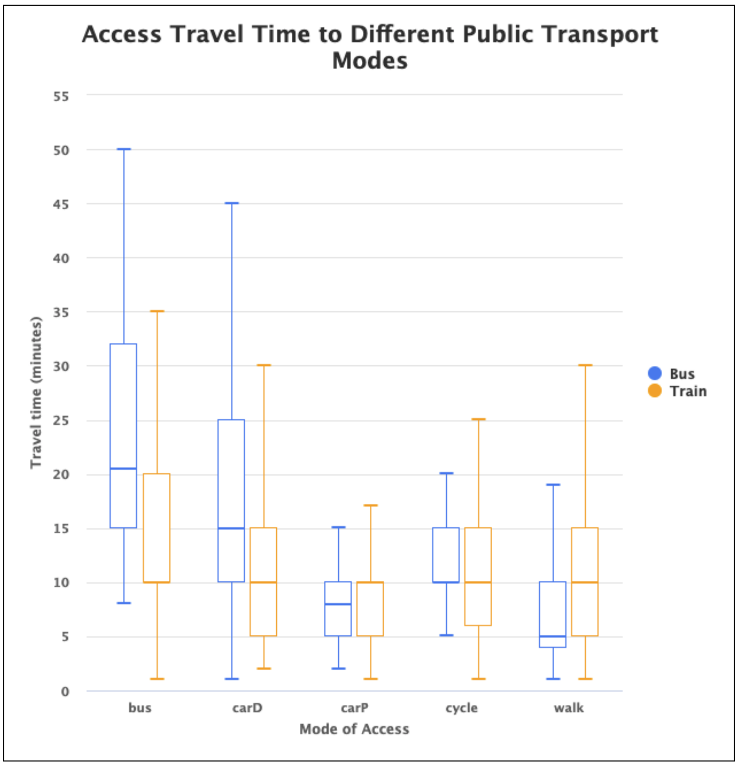

Figure 14. Access Travel Time to Different PT Modes

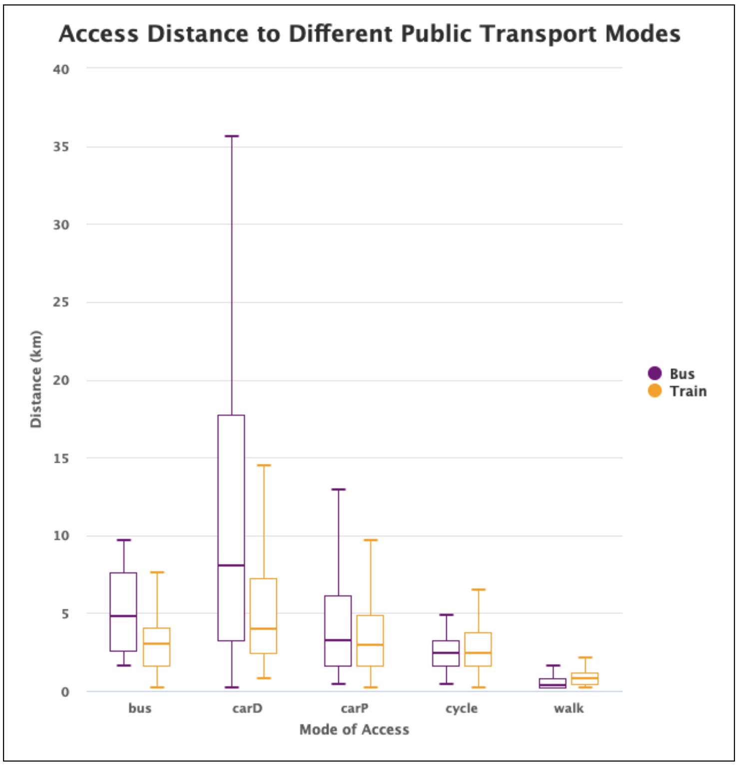

Figure 15. Access Distance to Different PT Modes

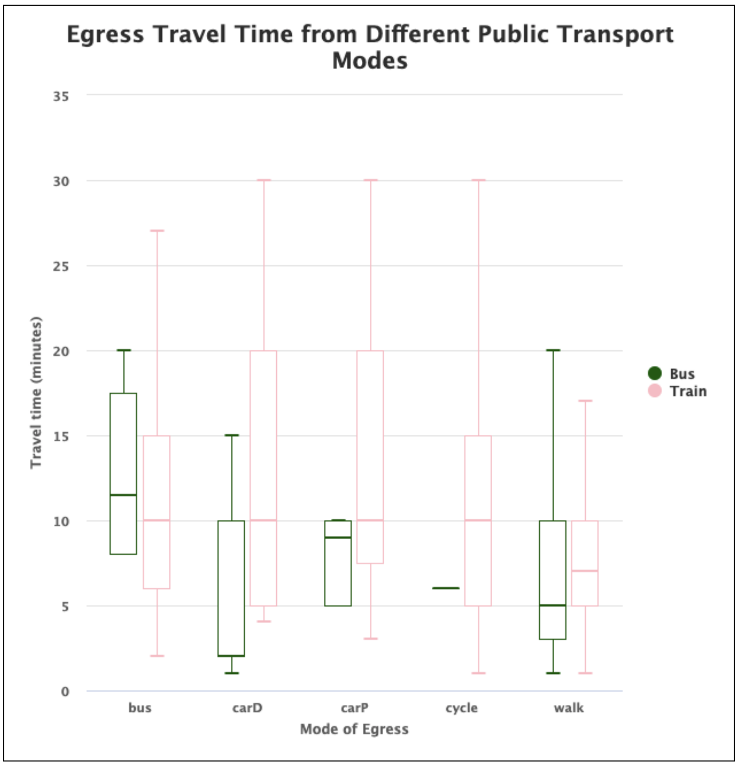

Figure 16. Egress Travel Time from Different PT Modes

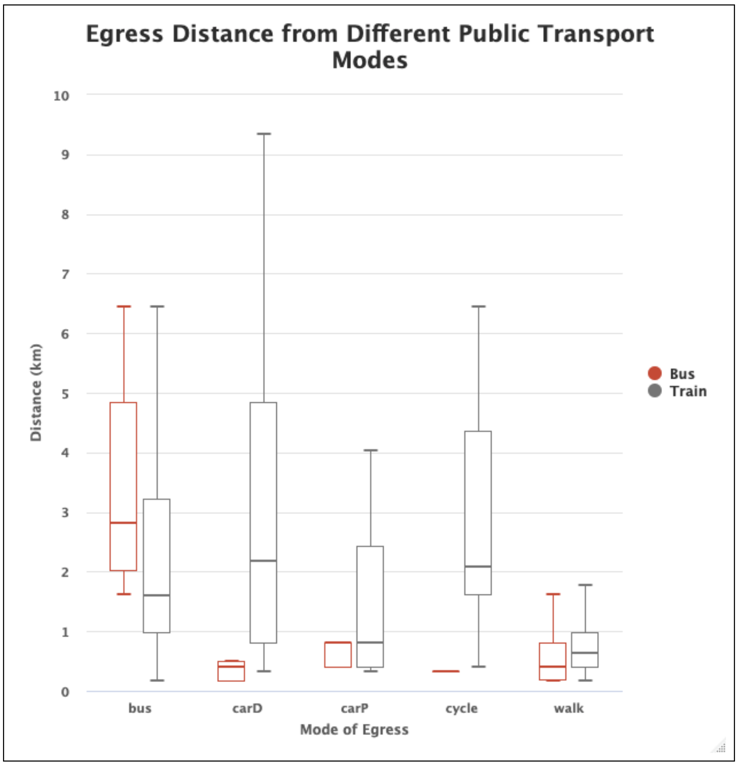

Figure 17. Egress Distance from Different PT Modes

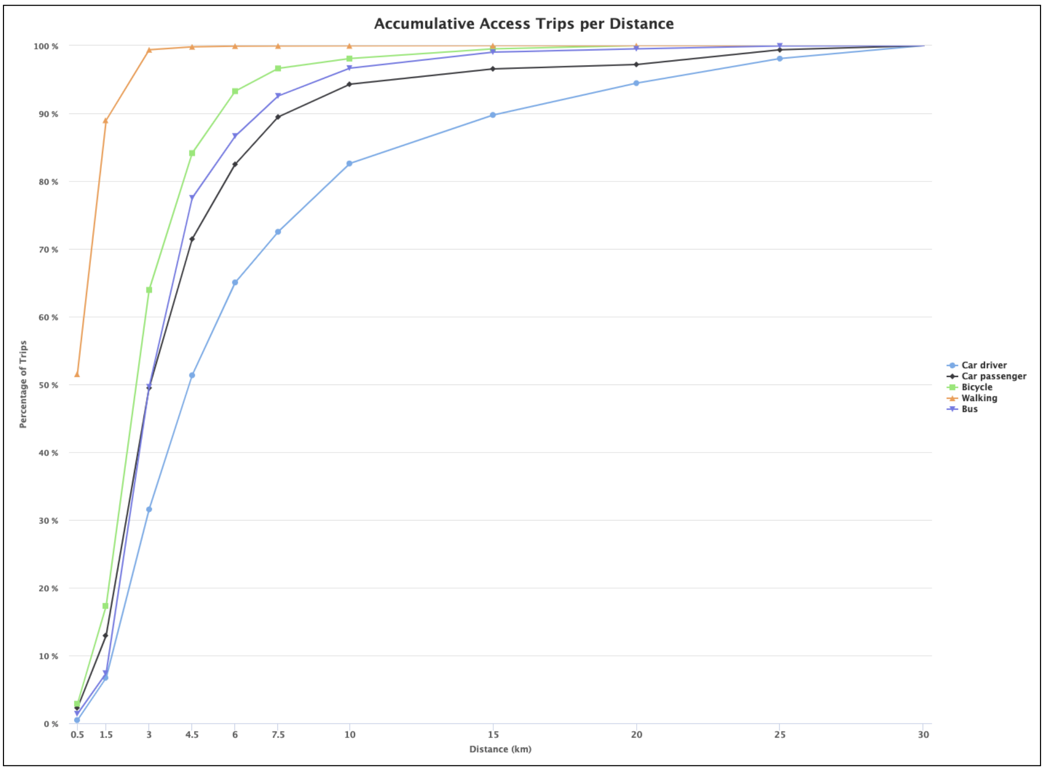

Figure 18. Accumulative Access Trips per Distance

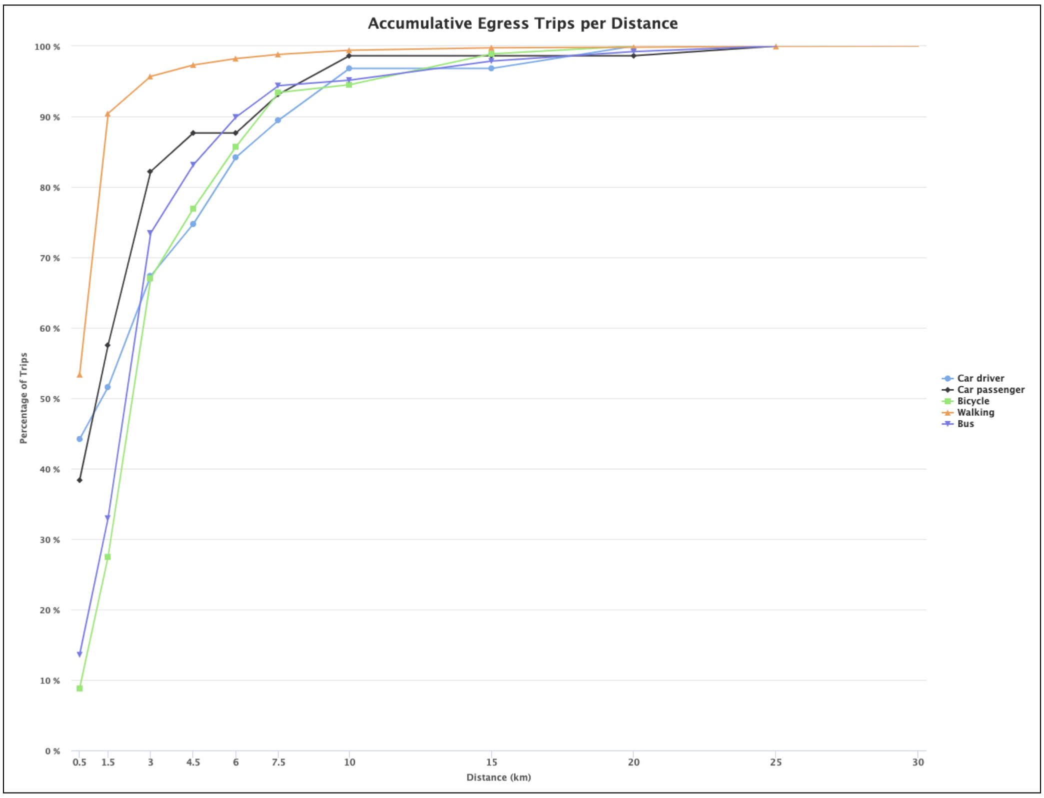

Figure 19. Accumulative Egress Trips per Distance

3.3. Modelling Methodology

The present thesis has opted for the utilization of the RStudio software in order to construct and assess statistical models pertaining to access and egress. RStudio represents an integrated development environment (IDE) based on open-source software that was designed to cater to the needs of the R programming language. This software platform is widely employed by researchers, data analysts, and statisticians as a means of performing data analysis, visualization, and modelling tasks.

Moreover, this research employs the Apollo package within the R programming language for the purpose of modelling the dataset. Apollo is a software package designed to facilitate the estimation and application of choice models within the R programming language. It provides users with a versatile set of tools, including pre-built model functions, as well as the ability to customize models to fit their specific research needs. Random heterogeneity, both at the level of individuals and observations and in both continuous and discrete forms, can be incorporated into all models. The package accommodates both standalone models and hybrid model structures and offers both classical and Bayesian estimation methods. Additionally, Apollo includes support for multi-threaded processing, which can expedite estimation procedures for large datasets. Moreover, the package is capable of analysing multiple discrete-continuous models, extending its application beyond standard discrete choice models [59, 60].

Multinomial & Nested Logit Models

This thesis aims to assess and compare the accuracy of the Multinomial Logit Model and Nested Logit Model frameworks in predicting access and egress. Both models will be utilized and evaluated to determine the optimal structure for predicting the outcome of interest. Multinomial logit (MNL) models and nested logit (NL) models are commonly used techniques in statistical modelling and econometrics for analysing and predicting categorical outcomes. [36, 38]

The multinomial logistic model posits that the data are case-specific, meaning that each independent variable has a unique value for every case. Additionally, the model assumes that the dependent variable cannot be predicted with complete accuracy from the independent variables for any given case. While independence between the independent variables is not a prerequisite, the model assumes a low degree of collinearity. Failure to satisfy this assumption makes it challenging to disentangle the distinct effects of multiple variables on the outcome of interest, thereby compromising the model's interpretability [61].

The MNL model assumes that the log odds of each potential outcome are linearly related to a typeset of prognosticator variables. Specifically, the chance of an outcome in a given category is sculptural as:

$$\Pr(Y_i = c) = \frac{e^{\beta_c \cdot \mathbf{X}i}}{\sum_{j=1}^{K} e^{\beta_j \cdot \mathbf{X}_i}} \tag{3.1}$$

Where $\Pr(Y_i = c)$ is the probability of the outcome being in category c, β_c is the vector of coefficients for the c$^{\text{th}}$ category, $\mathbf{X}$ is the vector of predictor variables, and $\sum_{j=1}^{K} e^{\beta_j \cdot \mathbf{X}}$ is the sum of the exponentiated values of the coefficients for all categories.

The nested logit model extends the application of logit modelling methods by enabling the incorporation of interdependence among responses, by organizing choices into distinct categories or "nests". Consequently, the observed outcome arises from a hierarchical decision-making process, with observed outcomes and attributes being linked to specific choices at each level. Furthermore, individual-specific characteristics may affect the decision-making process, independent of the outcomes under consideration [62].

The chance of a result in a given category is sculptured as:

$$\Pr(Y_{=}i) = \frac{\sum_j \left(e^{\frac{v_{ij}}{\lambda}} \theta_j\right)}{\sum_k \sum_j \left(e^{\frac{v_{kj}}{\lambda}} \theta_j\right)} \tag{3.2}$$

Where $\Pr(Y_{=}i)$ is the probability of the outcome being in category i, v_{ij} is the utility of alternative j in nest i, λ is a parameter that determines the correlation among alternatives in the Same nest, and θ_j is a nest-specific parameter that captures the correlation among alternatives in the same nest.

The main difference between MNL and NL models is that NL models allow for correlativity among alternatives in the same nest, while MNL models get into that each alternative is independent of the others. This makes NL models particularly useful for moulding situations where alternatives are not completely independent.

Correlation Matrices

Figure 20 presents the correlation matrix for access mode, highlighting significant associations between various variables. These include the number of cars and bicycles in a household, ownership of a driving license, household size, the number of adults, and their corresponding age groups, as well as household income category 3 (denoting the highest income category in the dataset). Moreover, the ownership of a bicycle exhibits a strong correlation with the number of bicycles in a household. Similarly, the number of children in a household is associated with their age group (1-19 years old) and escort activities..

Unemployment working status is found to be correlated with the number of adults in a household, the first two age categories (1-19 and 19-29) and the lowest income category. Additionally, retirement status is positively correlated with the highest two age categories (60+) and the lowest income category. Discretionary activities exhibit a negative correlation with the number of adults and cars in a household, and a positive association with the last two age categories (60+), whereas mandatory activities are positively correlated with the number of adults in a household, the number of cars and bicycles in a household, household size, and the highest income category. Mandatory activities also demonstrate a strong positive association with employment working status and the working age categories (2-3-4).

Access time and distance exhibit a strong positive correlation. Bus as the main travel mode is positively associated with the lowest income category, the youngest and oldest age groups, retirement working status, and discretionary activities. Conversely, it is negatively correlated with the number of cars in a household, ownership of a driving license, the highest income category, the third age category (30-49), total trip travel time, and mandatory activities. In contrast, train as the main travel mode exhibits the opposite correlation relationship with the same variables described in bus correlation relationships.

Car access mode, whether as a driver or passenger, exhibits similar correlations with other variables, primarily associated with the number of cars in a household, ownership of a driving license, and access distance and time. Walking is negatively correlated with the number of cars in a household, mandatory activities, access distance, and train as the main travel mode, whereas it shows a positive association with the bus as the main travel mode. Bus as an access mode is positively correlated with the train as the main travel mode and negatively correlated with the bus as the main travel mode. Finally, cycling appears to be noticeably correlated only with the ownership of a bicycle.

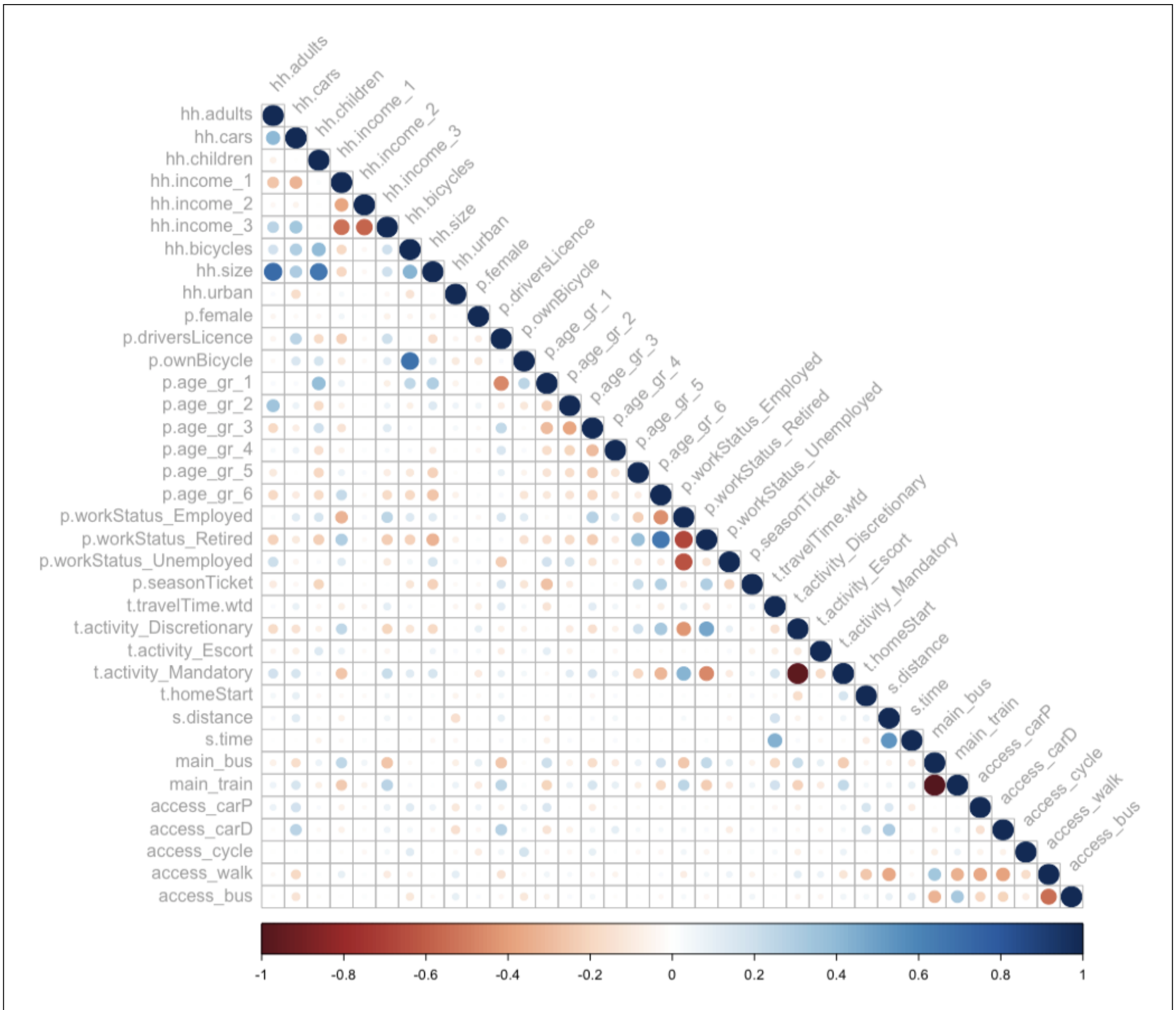

Figure 20. Access Correlation Matrix

The correlation matrix for egress mode is presented in Figure 21, indicates that the correlation relationships between various variables are quite similar to those observed in the access correlation matrix. However, there are some notable differences in the main mode associations with egress distance. Specifically, the matrix reveals that bus as a main travel mode exhibits a negative association with egress distance, whereas train as a main mode shows a positive association with the same variable. This finding suggests that individuals are less inclined to opt for the bus as their primary travel mode when the egress distance is long.

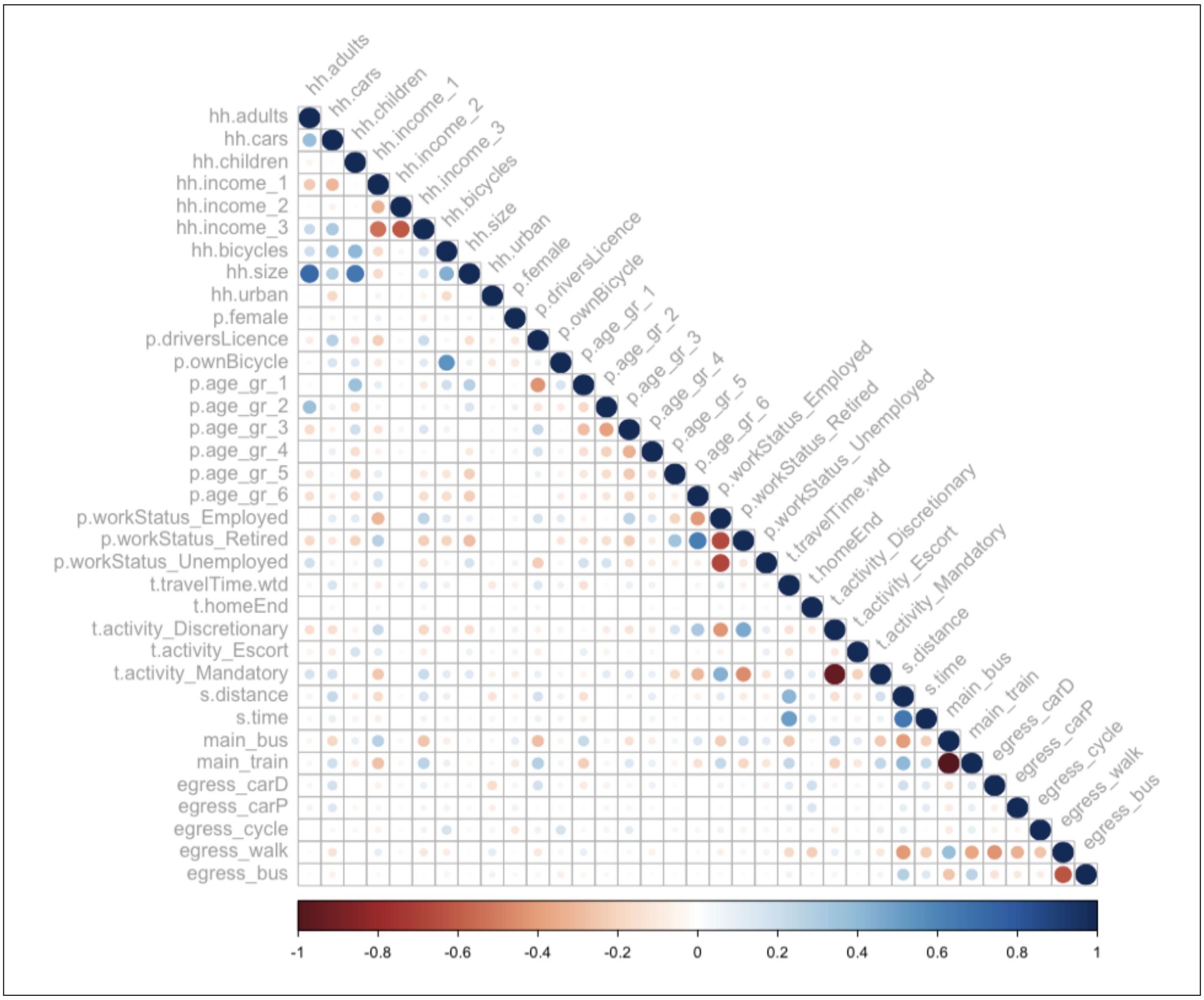

Figure 21. Egress Correlation Matrix

Nesting Structure

The determination of the nesting structure is contingent upon the modes that are believed to possess correlations in their unobserved utility components. During the process of model estimation, the degree of correlation can be delineated through the employment of the nesting parameter λ, whose value should range from 0 to 1 [36]. A value of $\lambda = 0$ would signify a complete correlation between the choices, indicating that they are essentially indistinguishable from one another. Conversely, a value of $\lambda = 1$ would denote a complete lack of correlation, resulting in the model becoming a standard logit. Any value outside this range typically deviates from the tenets of utility-maximizing behaviour, thus necessitating a re-estimation of the model.

Regarding the realm of forecasting, the nesting parameter functions as an indicator of the level of substitution that may arise between modes situated within the same nest, in comparison to those situated within other nests. A parameter that approximates 0 would suggest that the majority of substitutions occur within the nest, while a parameter that approximates 1 would imply that substitution is almost equally distributed. The potential ramifications of this issue are substantial in terms of promoting physical activity (PA). Specifically, the efficacy of any increases in cycling towards improving public health would be diminished if the newly adopted cycling behaviour were to supplant walking trips [63].

In order to evaluate potential correlations among active modes, two distinct nesting structures were tested and subsequently compared. The first nesting structure consisted solely of walk and cycle modes in the active nest and all motorized modes in the passive nest. The second nesting structure included bus, walk, and cycle modes in the active nest and car as a driver and a passenger in the passive nest. In conjunction with the active nest, the initial models presupposed the existence of a car nest, which consisted of both car drivers and passengers. The inclusion of these modes together in mode choice models is a common practice given their similarity [64]. However, if the nesting parameter λ for the car nest was deemed implausible, the car nest was eliminated, and the active nest models were re-estimated.

Model	McFadden R^2	$\lambda_{passive}$	λ_{active}	Active nest structure
Access Mode Choice	0.763	1.26	0.988	Walk + Cycle
	0.813	0.1881	0.5591	Bus + Walk + Cycle
Egress Mode Choice	0.701	0.998	0.436	Walk + Cycle
	0.734	0.3243	0.6888	Bus + Walk + Cycle

Table 9. Comparison of Mode Choice Nesting Structures

Table 9 displays the results of the initial model estimations inclusive of all coefficients. Removal of insignificant coefficients occurred only after the nesting structure was established and fixed for each model. This was done in order to prevent the inadvertent introduction of correlation to the error components during the removal process. While the nesting structure remained fixed, the estimation procedure allowed for variation in the nesting coefficients within a suitable range.

In terms of achieving a superior model fit outcome, the classification of buses as an access/egress mode in conjunction with the active nest exhibited more favourable results in comparison to alternative configurations. This was evidenced by a higher McFadden R^2 value and nesting coefficients that were found to be statistically significant in both models. The pertinent coefficients are visually distinguished in light grey shading within Table 9.

Reference Values

In the initial phase of the study, an assessment was carried out to establish the reference value for each variable that comprises multiple categories. The assessment was based on the frequency of instances recorded within each category, which led to the identification of the following reference points: Age, which was designated as the third age category (30-49); work status, which was determined to be employed; activity type, where mandatory activities were selected as the reference category; and main travel mode, which was identified as train. However, only one variable showed an exception, namely income, where the base value was set to the highest income category. This decision was made due to the correlation between the lowest and highest income categories and other variables. For variables with binary values, the reference variable was assigned to False, with the exception of gender, for which the base reference was assigned to males.

The modelling process involved conducting an examination of the correlation matrix between the variables and access/egress modes individually to attain a more comprehensive understanding of their interaction. Following this, each attribute was individually tested by running the model with a singular attribute and primary travel mode to comprehend the interaction within the model and achieve a deeper comprehension of the correlation. In a later stage, this process was replicated for the combined primary travel modes.

3.3.1. Modelling Access

Two distinct modelling frames, namely Multinomial Logit (MNL) and Nested Logit (NL), were employed to model access data and compare the performance of the two models in determining the optimal approach for modelling access data. Additionally, two discrete models were developed with the specific objective of isolating the primary modes of transportation, namely bus and train, in order to investigate their individual outcomes.

NL Model

The statistical model[1]has been fitted and evaluated using a regression analysis, which is used to describe the relationship between a dependent variable and one or more independent variables. The model returned the following key statistics:

Rho square adjusted (McFadden) = 0.80

This is a measure of the goodness of fit of the model, adjusted for the number of parameters and the sample size. An adjusted R-squared of 0.73 indicates that the model explains 73 per cent of the alternative in the outcome variable after adjusting for the number of parameters in the model and the sample size.

Log-likelihood = -12239.74

This is the maximum value of the likelihood function for the model, which is a measure of how well the model fits the data. A higher log likelihood indicates a better fit between the model and the data. In this case, the log-likelihood is -15908, which suggests that the model fits the data reasonably well.

Akaike Information Criterion (AIC) = 24591.48

This is a measure of the relative quality of the model, taking into account both the goodness of fit and the complexity of the model. The AIC penalizes models with more parameters, so a lower AIC

[1]The regression summary table is located within the appendix section (Table A3).

indicates a better trade-off between model complexity and goodness of fit. In this case, the AIC is 31937, which suggests that the model is relatively good.

Bayesian Information Criterion (BIC) = 25014.19

This is another measure of the relative quality of the model, similar to the AIC but with a stronger penalty for more parameters. A lower BIC indicates a better trade-off between model complexity and goodness of fit. In this case, the BIC is 32367, which is higher than the AIC and suggests that the model is relatively complex.

Comparing Access NL and MNL Models

When comparing the key statistics of the NL and MNL models, the NL model demonstrates superior performance. The NL model has a higher R2 adjusted value ***(0.80 vs. 0.73)*** indicating better explanatory power, a higher log-likelihood value ***(-12239.74 vs. -14983.65)*** indicating a higher likelihood of producing observed data, and lower AIC ***(24591.48 vs. 26465.3)*** and BIC ***(25014.19 vs. 26935.18)*** values indicating a better overall fit for the data. Therefore, based on these key statistics, the NL model outperforms the MNL model.

Access NL Model (Bus)

Upon exclusively designating the bus as the primary mode of travel, the model[2]generated the subsequent key statistical metrics.

- R squared adjusted: 0.92
- Log-likelihood: -2397.438
- AIC: 4904.876
- BIC: 5259.742

Access NL Model (Train)

Upon exclusively designating the train as the primary mode of travel, the model[3] generated the subsequent key statistical metrics.

- R squared adjusted: 0.69
- Log-likelihood: -9601.793
- AIC: 19313.59
- BIC: 19706.38

3.3.2. Modelling Egress

Similar to the access model, two distinct modelling frames, ***MNL*** and ***NL*** , were employed to model egress data and compare the performance of the two models in determining the optimal approach for modelling egress data. Additionally, two discrete models were developed with the specific objective of isolating the primary modes of transportation, namely bus and train, in order to investigate their individual outcomes.

[2]The regression summary table is located within the appendix section (Table A1).

[3]The regression summary table is located within the appendix section (Table A2).

NL Model

The statistical model[4] has been fitted and evaluated using a regression analysis, which is used to describe the relationship between a dependent variable and one or more independent variables. The model returned the following key statistics:

The adjusted R2 value of ***(0.73)*** suggests that the model explains 73 per cent of the variance in the dependent variable after adjusting for the number of independent variables used in the model. The log-likelihood value of ***(-15908.63)*** is a measure of how well the model fits the data, with lower values indicating a better fit. The AIC value of ***(31937.26)*** is a measure of the model's quality and simplicity, with lower values indicating a better balance between the goodness of fit and model complexity. The BIC value of ***(32367.68)*** is another measure of model quality and simplicity, similar to AIC but with a stronger penalty for model complexity.

Comparing Egress NL and MNL models

The NL model outperforms the MNL model in several key statistics. The NL model has a higher R-squared adjusted value ***(0.73 vs. 0.69)*** indicating better explanatory power, a higher log-likelihood value ***(-15908.63 vs. -18070.16)*** indicating a higher likelihood of producing observed data, lower AIC ***(31937.26 vs. 33250.32)*** and BIC ***(32367.68 vs. 35644.87)*** values indicating a better overall fit and parsimony. Consequently, the NL model is a better choice than the MNL model for the given data.

Egress NL Model (Bus)

Upon exclusively designating the bus as the primary mode of travel, the model[5] generated the subsequent key statistical metrics.

- R squared adjusted: 0.92
- Log-likelihood: -2397.438
- AIC: 4904.876
- BIC: 5259.742

Egress NL Model (Train)

Upon exclusively designating the train as the primary mode of travel, the model[6] generated the subsequent key statistical metrics.

- R squared adjusted: 0.69
- Log-likelihood: -9601.793
- AIC: 19313.59
- BIC: 19706.38

[4] The regression summary table is located within the appendix section (Table A6).
[5] The regression summary table is located within the appendix section (Table A4).
[6] The regression summary table is located within the appendix section (Table A5).

4. Results & Discussion

This chapter presents the outcomes of the statistical models employed in this study, which aimed to explore the relationship between different attributes and their impact on access/egress mode choice. In this chapter, the results of the statistical models and their implications will be discussed. Firstly, the outcomes concerning individual main travel modes, namely bus and train will be presented followed by the results obtained from the combined model for access/egress separately. The analysis presented herein is intended to provide insights into the complex interplay of different factors and their impact on the outcome of interest, with the ultimate goal of informing future research and practical applications.

Graphs were constructed using data points that met a statistical significance criterion of p-value ≤ 0.1

4.1. Access Models

4.1.1. Bus Access Model

According to the model outcomes (Figure 22), residing in urban areas has a substantial positive impact on the selection of bus and walking as access modes to the bus, possibly owing to the availability of buses, good infrastructure that encourages walking, or the wider distribution of bus stations in urban areas. Additionally, the number of cars in a household significantly affects mode choice, with households owning a car or more showing a negative inclination towards choosing any other mode for access. This could be due to the comfort and convenience cars offer compared to other modes.

Age also plays a crucial role in selecting access modes, with biking as an access mode being generally unpopular across all age groups, possibly due to the limited storage space for bikes on buses. Individuals under the age of 18 are more likely to be driven to the bus station, likely because many parents drive their children to school or transit stations. Those between 19 and 29 years are more likely to take the bus or walk to the bus station and slightly inclined towards taking the car as passengers, while older groups prefer not to walk or bike as their access modes to bus stations.

Income significantly influences people's choices, with lower or middle income are more inclined to avoid travelling as car passengers. This can be attributed to the fact that lower-income households are less likely to possess a car, while individuals in middle-income families often have access to their own vehicle. Additionally, lower-income individuals are more inclined to walk to the bus station, in contrast to their middle-income counterparts. Females are less likely to use bikes or walk to the bus station and are slightly more inclined towards being driven.

Owning a driving license influences individuals' mode choice as they are more likely to drive a car if they have access to one. Those who own a bike are more inclined to use it as their access mode. Season ticket ownership influences people's choices negatively, as they are less likely to take the bus in favour of travelling as a car passenger, which may be because not all season tickets are valid on buses in England.

Unemployment appears to slightly increase bike usage, as individuals try to cut down on expenses,

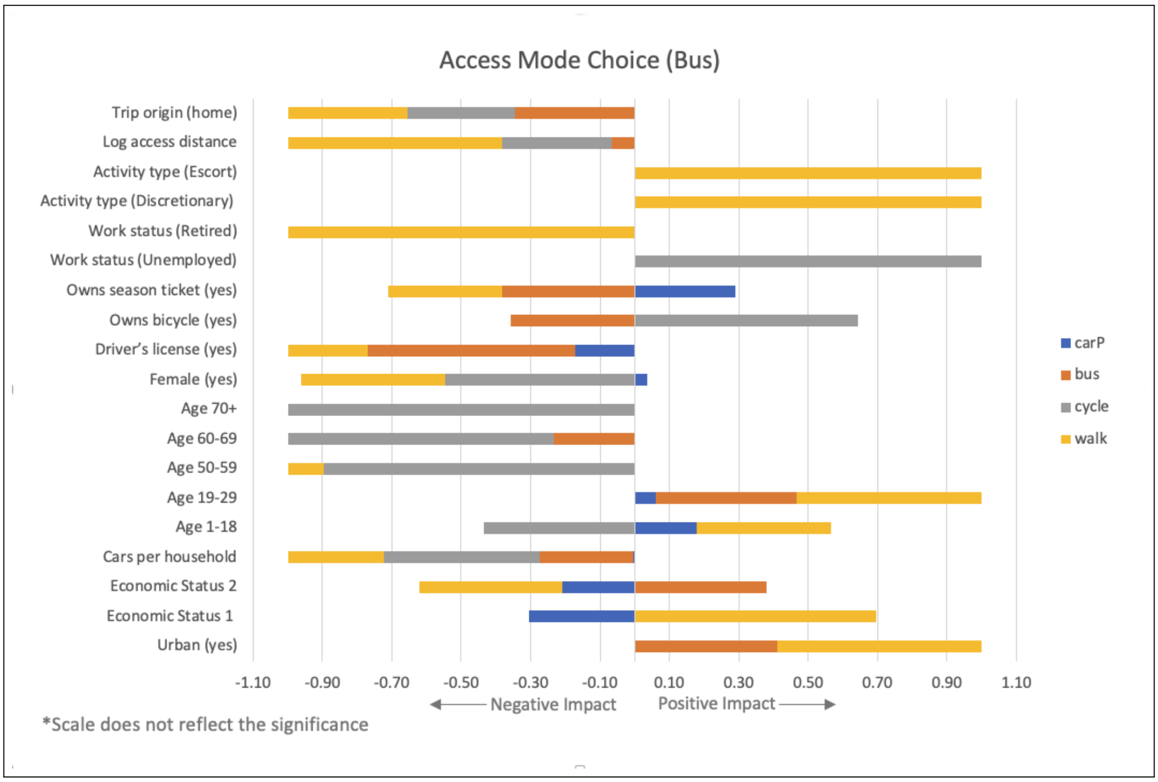

Figure 22. Access Mode Choice (Bus)

while retired individuals do not prefer to make their access trips on foot. When the bus is the main travel mode for discretionary or escort purposes, walking seems to be the preferred mode of access rather than other modes.

Lastly, longer access distances negatively impact active modes, with walking being the most affected, followed by cycling and taking the bus. This is reasonable, considering that access trip lengths are generally longer for motorized modes compared to active modes. Having the trip origin at home appears to incentivize car usage as a driver or passenger over taking other modes.

4.1.2. Train Access Model

Based on the model outcomes (Figure 23), it is evident that urban residency has a significant positive impact on the selection of bus and walking as access modes to the train. This can be attributed to the availability of buses, wider distribution of bus stations, consistent bus schedules, and good infrastructure that encourages walking in urban areas. Additionally, the number of cars owned by a household significantly affects mode choice, with households possessing one or more cars showing a negative inclination towards choosing any other mode for access. This may be due to the comfort and convenience cars offer, as well as the availability of parking lots near train stations which is not the case with bus stations.

The age of individuals also plays a crucial role in selecting access modes. Biking as an access mode is generally unpopular across almost all age groups included in the analysis, potentially due to factors such as the lack of dedicated infrastructure, extra fees for bringing bikes on board trains, or weather conditions across most of England throughout the year. Individuals under the age of 18 are more likely to be driven to the train station, likely due to parents driving their children to school or transit stations. Those between 19 and 29 years are more likely to take the bus, walk, or travel in

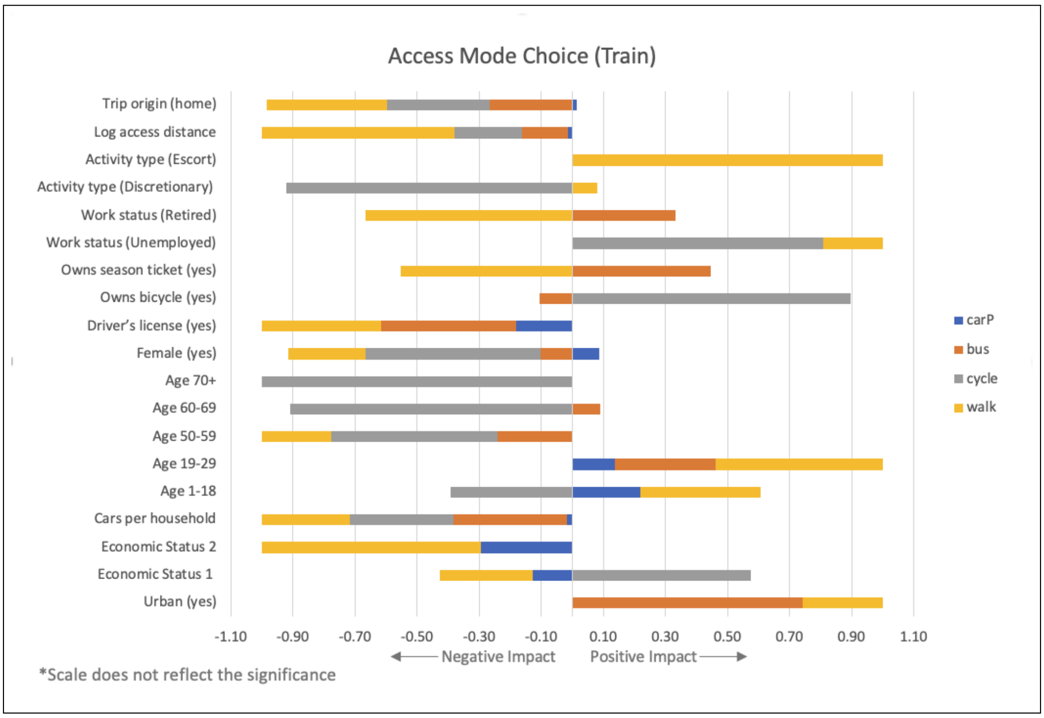

Figure 23. Access Mode Choice (Train)

a car as passengers to the train station, while older groups prefer not to walk for their access trips.

Income significantly influences people's choices, with lower or middle-income households more inclined to avoid travelling as car passengers. This can be attributed to the fact that lower-income households are less likely to own a car, while individuals in middle-income families often have access to their own vehicle. People belonging to the two aforementioned categories tend to not make their access trips on foot, and people belonging to lower-income households seem to use the bike more than other modes.

Females are less likely to use the bus, bikes, or walk to the bus station and are more inclined towards being driven. Owning a driving license influences individuals' mode choice as they are more likely to drive a car if they have access to one. Those who own a bike are more inclined to use it as their access mode. Season ticket ownership does not have much of an impact except for discouraging walking to the train station. Unemployment appears to slightly increase active travel (bike and walk), as individuals try to cut down on expenses, while retired individuals prefer not to make their access trips on foot and prefer the bus instead. When the train is the main travel mode for discretionary or escort purposes, walking seems to be the preferred mode of access rather than other modes.

Lastly, longer access distances negatively impact active modes, with walking being the most affected, followed by cycling and taking the bus. This is reasonable considering that access trip lengths are generally longer for motorized modes compared to active modes. Having the trip originate at home appears to incentivize car usage as a driver or passenger over taking other modes.

4.1.3. Access Model (Bus & Train)

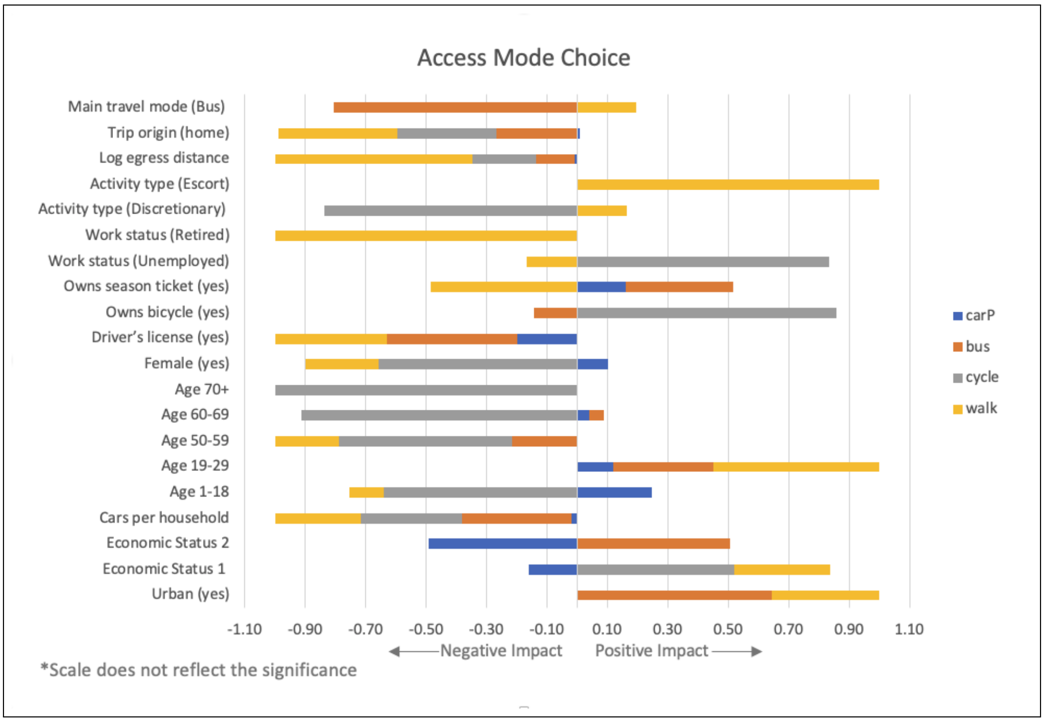

Figure 24. Access Mode Choice

The results of the model (Figure 24) indicate that the use of buses as the main travel mode discourages individuals from using them as an access mode, instead favouring walking. In urban areas, the availability of buses, their wide distribution, consistent schedules, and the prevalence of pedestrian-friendly infrastructure incentivize people to use buses and walk as access modes. However, there is no evidence to suggest that these factors affect the choice of using bikes as an access mode.

The possession of cars by households significantly influences mode choice, with households owning one or more cars exhibiting a negative inclination towards selecting any other mode of transportation for access. This preference may be attributed to the comfort and convenience that cars offer. Age is another important factor in the selection of access modes, with biking being an unpopular option across all age groups. The limited storage space for bikes on buses, the lack of dedicated infrastructure, extra fees for bringing bikes on board trains, or weather conditions in most parts of England throughout the year may contribute to this preference. Individuals under 18 years of age are more likely to walk or be driven to the bus or train station, likely because parents often drive their children to school or transit stations. Those aged 19 to 29 years are more likely to take the car as passengers, the bus, or walk to the transit station, while older age groups prefer not to walk or bike for their access trips. Individuals over 60 years are slightly more inclined to make their access trips as car passengers.

Income plays a significant role in determining people's mode choice for access trips. Lower or middle-income households are more likely to avoid travelling as car passengers, which can be attributed to their lower likelihood of owning a car. On the other hand, individuals in middle-income families often have access to their own vehicles. Moreover, people belonging to middle-income families tend to use the bus as their preferred access mode, while lower-income family members are more inclined to use the bike and less likely to choose walking. The preference for not walking may be due to access trips to train stations, taking into consideration the sample size of people using the train as their main travel mode and the preference for other access modes when going to train stations.

Females are less likely to use bikes or walk to the bus station and are more inclined towards being driven. The possession of a driving license negatively influences individuals' mode choices regarding active travel modes and the bus, as they are more likely to drive a car if they have access to one. Similarly, owning a bike encourages its use as an access mode in favour of other modes. Unemployment appears to increase bike usage, while retired individuals do not prefer to make their access trips on foot. Discretionary activities negatively affect bike usage while slightly encouraging walking, while escort activities seem to be only positively associated with walking.

Finally, longer access distances negatively impact active modes and the bus, with walking being the most affected, followed by cycling and taking the bus. Having the trip origin at home appears to incentivize car usage as a driver or passenger over taking other modes.

4.2. Egress Models

4.2.1. Bus Egress Model

The use of bicycles as an egress mode from buses is found to be almost non-existent, as demonstrated by Figures 16 and 17. To address this, the ownership of bicycles was excluded from the bus egress model during estimation, and associations with cycling as an egress mode were removed from most of the attributes.

The results of the model (Figure 25) indicate that residing in urban areas has a significantly positive effect on the selection of walking as an egress mode from buses. This may be due to the relatively short distances from bus stations to final destinations in urban areas as shown in Figure 17. Moreover, the number of cars owned by a household has a significant impact on mode choice, with households owning one or more cars showing a negative inclination towards choosing any other mode of egress.

Age is also a critical factor in selecting egress modes. Individuals below the age of 18 are more likely to walk or be driven from the bus station, while those aged between 19 and 29 years are more likely to take the bus or walk from the bus station. On the other hand, older age groups prefer not to make their egress trips on foot and opt for motorized modes, particularly the bus.

Household size seems to affect the prioritization of travelling as passengers or walking over taking the bus. Being female reduces the likelihood of walking or cycling from the bus station. Owning a driving license has a significant influence on individuals' mode choices as they are less likely to walk or use the car as passengers, and season ticket ownership also affects people's choices in walking from bus stations.

Unemployment appears to increase bus and car usage as passengers, while retired individuals prefer to use the bus or make their egress trips on foot, which may indicate that retired people take the bus when available and prefer to walk for very short trips. When the bus is the primary travel mode, discretionary activities encourage car usage as passengers, while escort activities promote walking as the preferred egress mode. Finally, longer egress distances have a negative impact on active modes and having the trip destination at home appears to incentivize car usage rather than taking the bus or walking.

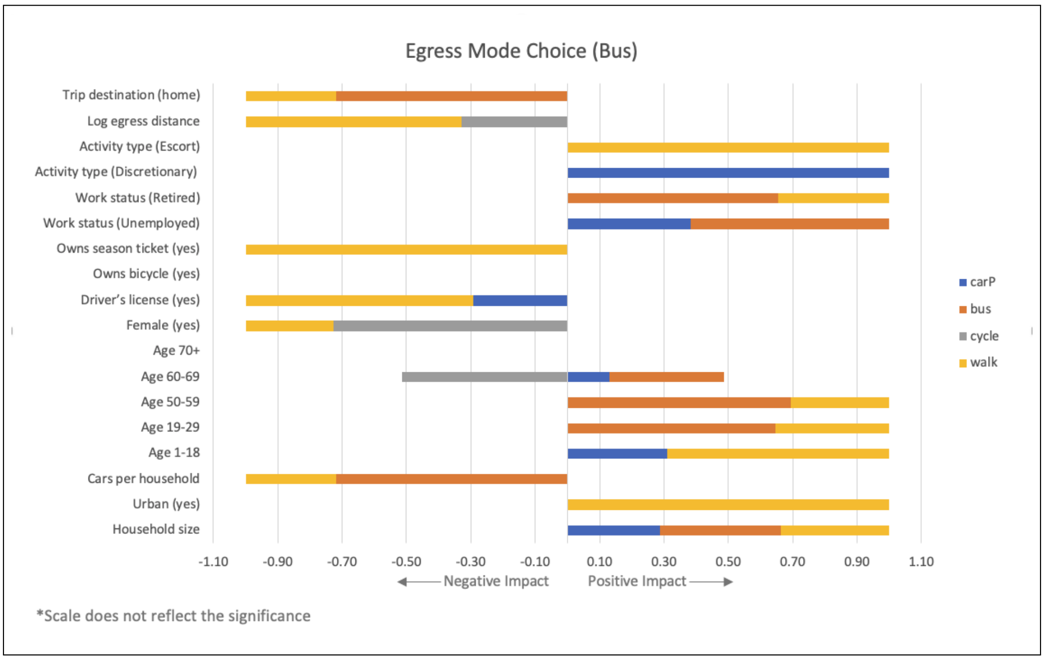

Figure 25. Egress Mode Choice (Bus)

4.2.2. Train Egress Model

The findings of the model (Figure 26) suggest that urban residence has a significant positive influence on the selection of bus, cycling, and walking as egress modes from train stations. This could be attributed to the reliability of buses in urban areas and the relatively short distances between train stations and final destinations in urban areas, as demonstrated in Figure 17. In addition, the number of cars owned by a household has a considerable impact on mode choice, with households possessing one or more cars showing a negative preference for any other egress mode.

Age is also a crucial factor in choosing egress modes. Individuals below 18 years of age are receptive to all egress modes, as are those aged between 19 and 29 years. In contrast, older age groups prefer motorized modes over active modes.

Household size appears to influence the priority placed on using the bus, bikes, or walking for egress. Being female reduces the likelihood of walking or cycling from the train station and increases the chance of using the car as a passenger. Owning a driver's license has a significant impact on individuals' mode choices, making them less inclined to walk or use the bus. Possessing a bike increases the chances of using it as an egress mode, while season ticket ownership has a positive effect on selecting the bus for egress trips.

Unemployment is seen to decrease bus usage and walking trips, while retired individuals prefer taking the bus for their egress trips. Discretionary activities have a negative impact on cycling during egress trips, while escort activities seem to be slightly inclined to not use the car as a passenger

Lastly, longer egress distances negatively affect active modes and travelling as a passenger and promote the use of buses, and having the trip destination at home appears to encourage car usage over taking the bus, cycling, or walking.

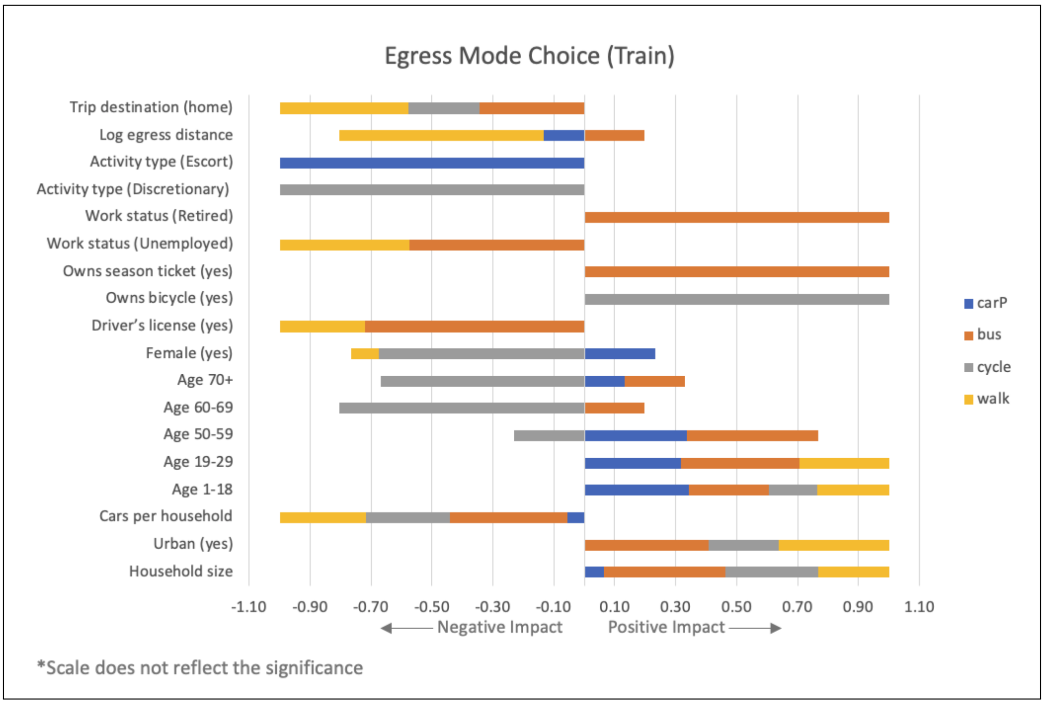

Figure 26. Egress Mode Choice (Train)

4.2.3. Egress Model (Bus & Train)

The results of the model (Figure 27) indicate that urban living significantly increases the likelihood of selecting the bus, cycling, and walking as egress modes. This could be due to the reliability of buses in urban areas and the relatively short distances between transit stations and final destinations, as illustrated in Figure 17. Additionally, the number of cars owned by a household has a considerable negative impact on the preference for any other egress mode.

Age is a critical factor in determining egress mode choice. Individuals under the age of 18 are receptive to all egress modes, except cycling, which shows no significant evidence. This is reasonable given that a vast majority of this age group's trips are to and from school. Conversely, those aged 19-29 years are less likely to walk, favouring using the bus or car as passengers. In contrast, older age groups prefer motorized modes over active modes, given their higher likelihood of having access to a car and a physical condition that discourages them from taking active modes.

Household size appears to influence the priority placed on using the car as a passenger, bus, bikes, or walking for egress. This might be due to households with more people having a higher probability of containing children, which explains the preference for all modes of egress. Being female reduces the likelihood of choosing the bus, cycling, or walking and increases the chance of using the car as a passenger. Owning a driver's license has a significant impact on individuals' mode choices, making them less inclined to use other modes of egress in favour of using the car as drivers. Possessing a bike increases the chances of using it as an egress mode, while season ticket ownership has a positive effect on selecting the bus for egress trips. Unemployment is associated with using the car as a passenger, which is likely related to people under 18 who do not work, while retired individuals prefer taking the bus for their egress trips. Discretionary activities have a negative impact on using the bus, cycling or walking during egress trips, while escort activities seem inclined to not use the car as a passenger and slightly prefer walking.

Finally, longer egress distances negatively affect active modes and travelling as a car passenger

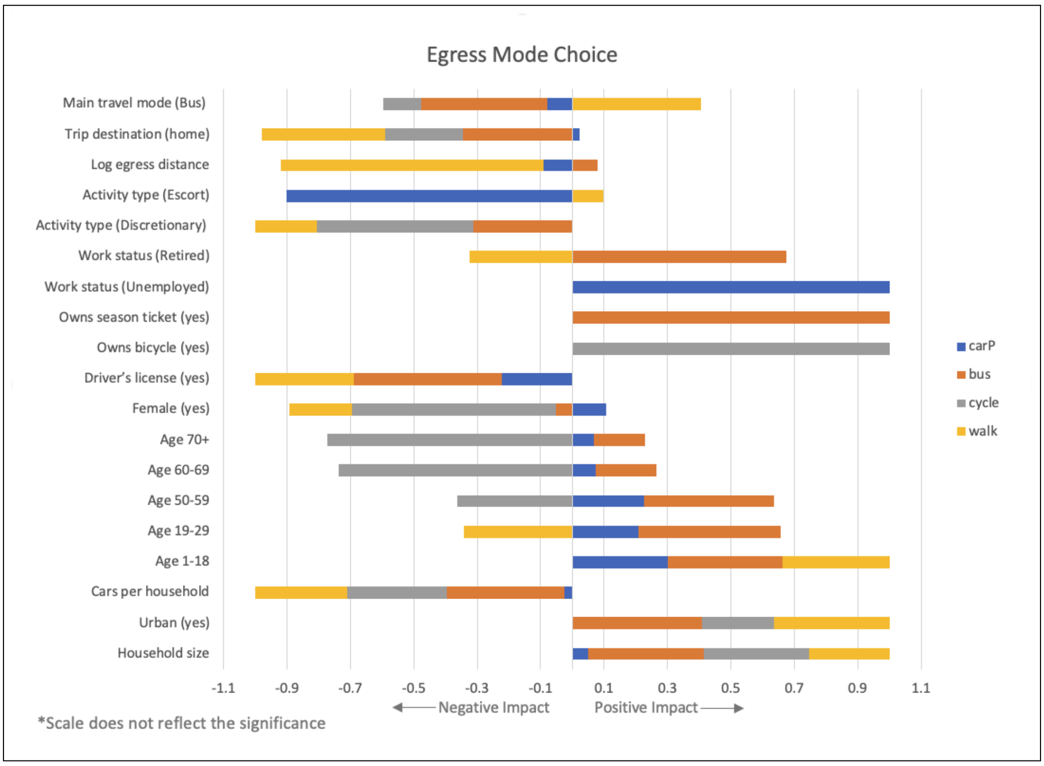

Figure 27. Egress Mode Choice

and promote the use of buses. This observation may be due to the unavailability of the car at the egress end and the unlikelihood of walking or cycling for longer distances. Having the trip destination at home encourages car usage over taking the bus, cycling, or walking while travelling in the bus as the primary travel mode decreases the chance of using all egress modes except for walking.

4.3. Discussion

The significance of attributes found in the access mode does not necessarily carry over to the egress mode. Several factors may account for this disparity, including variations in origin and destination locations, such as differences in distance or quality of pedestrian infrastructure. Individuals may be willing to walk long distances to access public transportation if the walking path is pleasant and safe but may be less willing to do so if it is poorly lit or located in an unsafe area.

Similarly, the characteristics of the public transport mode itself may differ, such as in frequency, speed, and capacity. For example, individuals may be more willing to use a bus for egress mode if it offers a direct and frequent route but may be less willing to use it for access mode if it has a low frequency or is overcrowded, particularly when travelling from suburban to urban or business areas.

Individual preferences and behaviours may also play a role. While some people may be willing to walk a certain distance to access public transport, they may prefer to take a taxi or a bus for egress mode due to personal preferences or time constraints.

Notably, the economic status was found to have no significant effect on the egress mode and a minimal impact on the access mode when using the bus as the primary mode of travel. This may be attributed to the lack of a cost variable in the estimation, which would have provided greater clarity and depth to the economic status. The author decided not to include a cost factor in the model due to a lack of sufficient evidence on people's perceptions of cost in access/egress. For instance, would travellers pay more for the access trip, such as a £20 taxi fare to get to the train station to take a £5 train ride for comfort reasons? Or, would they, alternatively, upgrade their class ticket? How would these scenarios change if the traveller was limited on time? Would they respond by paying a higher cost for the access trip or changing their primary mode of travel altogether?

There is a substantial body of literature examining the costs associated with the main mode of travel, but very little research has focused on access/egress costs. Without a comprehensive understanding of these costs and the evidence required to simplify the complexity of access/egress cost allocation decisions, it may not be the best approach to include a cost variable in the model.

Another possible explanation for the reduced significance of economic status in comparison to conventional travel mode choice models is that decisions are influenced by the availability of access/egress modes, which can vary considerably depending on the area/city. For instance, a traveller may be willing to pay more for a scooter or shared mobility option for access/egress purposes. However, this aspect could not be examined thoroughly due to the scarcity of available records and the documentation system, which combined the trips into a single mode and was restricted to the London area exclusively. Furthermore, the dataset provided no information that could be used to link the trips to specific zones or areas.

In terms of active modes for access/egress, females have been observed to prefer motorized modes. This tendency may be attributed to various factors, such as perceiving walking or cycling as less secure than travelling in a car, particularly in areas with deficient infrastructure or high levels of crime. Furthermore, women may face time constraints and responsibilities, such as looking after children or other family members, which could make walking or cycling impractical. Nevertheless, it is noteworthy that these factors are multifaceted and may vary according to the individual's specific circumstances and the context in which they occur.

5. Conclusion

This study employs a discrete choice model to investigate travel mode choice for access and egress. To achieve this, the National Travel Survey dataset was utilized, which provides comprehensive information on the travel behaviours of residents of England within Great Britain. The model framework incorporates socio-demographic, spatial, and trip characteristic indicators. Furthermore, different model types were compared to identify the most appropriate approach to modelling the available dataset. Lastly, access and egress travel modes were modelled separately, with each having its variations.

The findings of this study largely validate the conclusions drawn in previous research. Factors such as distance to transit stations, car availability, gender, and age group were the most influential indicators in all the models. Moreover, the primary travel mode significantly influenced the choice of access and egress modes. Work status and activity type demonstrated a significant impact in certain situations, while it had a lesser effect in others. Furthermore, ownership of a bicycle, driving licences, and transit passes had a significant impact but only in select situations, and they were useful in interpreting the outcomes of the models.

5.1. Limitations

The study examined several variables that store information on satisfaction with walking and cycling infrastructure but found that they returned significantly fewer records than required for reliable modelling. Moreover, their results contradicted those of other variables. A correlation matrix analysis revealed a negative correlation between satisfaction with walking and cycling infrastructure and the use of walking or biking as access/egress methods. The initial model runs confirmed this finding. As a result, these two variables were excluded from the estimation, which limited the ability to -partially- test the impact of environmental beauty on access/egress mode choice.

One plausible explanation for this outcome may be that although affluent areas or neighbourhoods typically have a better surrounding infrastructure, residents of such areas are more likely to use cars as passengers or drivers, which could classify many of them as captive car users. Although scientific evidence for a correlation between affluence and better walking infrastructure was not found, it is a widely accepted observation. This could be due to the greater political influence and access to public funding that affluent areas or neighbourhoods often have, as well as their historical development.

The variable pertaining to household area classification was excluded from the estimation due to its inconsistent results, as well as the inadequate number of records for certain access/egress modes that hindered reliable modelling. This variable introduced a classification system that combined urbanity and economic classification into eight categories. However, several categories had insufficient records for certain access/egress modes to be modelled, and merging categories was not feasible for two reasons. Firstly, some categories with similar classifications exhibited different correlation signs with other variables. Secondly, categories with similar classification and correlation relationships lacked sufficient records even after being combined. The exclusion of this variable limited the ability to capture the impact of environmental beauty on access/egress mode choice in another way that was initially planned.

5.2. Broader Issues

The available dataset was inadequate for evaluating the impact of contextual factors and environmental beauty on travel mode choice, as it lacked variables that could have provided insights into weather conditions, travel companions, reliability of access/egress modes, geographic location, and quality of surrounding infrastructure. Furthermore, important variables, such as the safety index of the neighbourhood and the presence of health conditions that limit mode choice, were absent from the dataset. Additionally, certain travel modes, including shared mobility, carpooling, and scooters, were not adequately represented in the dataset.

If the variables outlined in Table 3 (highlighted in light grey) and explained in Chapter 3 (3.2.2. Harmonising and Preparing the Data) had been properly captured, they would have provided indispensable insights into the impact of area classification on travel mode choice. Moreover, the selected time period's suitability remains an ambiguous proposition. The study focused on a 7-day period, which failed to consider seasonal variation in travel behaviour. For instance, El-Assi et al.'s [65] research discovered significant seasonal differences in active travel modes, attributable to weather-related impacts [63].

5.3. Final Remarks

This thesis incorporates separate models for access and egress, providing a more detailed analysis as these modes possess unique characteristics and mode choice behaviours. Furthermore, separate modelling allows for a better understanding of the attributes that influence access/egress mode choice. However, this approach is simplistic and carries both advantages and disadvantages.

A more detailed approach would be to model the two modes jointly or even model each mode jointly with the main travel mode. The former approach would be beneficial in obtaining a more accurate prediction and a comprehensive understanding of the entire trip from door to door. For instance, if an individual chooses a certain travel mode for easier access, this choice could be taken into consideration when predicting their egress mode, which would eventually lead to better interaction between modes. Joint modelling of access/egress with the main travel mode could also offer a holistic analysis of the entire travel system, including the potential for multimodal trips and the impact of different modes on travel behaviour and sustainability. Additionally, joint modelling enables a better comprehension of the intricate interactions and trade-offs between different modes, as well as the factors influencing mode choice and travel behaviour.

6. Bibliography

[1] C. Nobis, “Multimodale vielfalt. quantitative analyse multimodalen verkehrshandelns,” *Humboldt-University of Berlin*, 2014.

[2] L. Meyer de Freitas, H. Becker, M. Zimmermann, and K. W. Axhausen, “Modelling intermodal travel in switzerland: A recursive logit approach,” *Transportation Research Part A: Policy and Practice*, vol. 119, pp. 200–213, 2019.

[3] T. Wörle, L. Briem, M. Heilig, M. Kagerbauer, and P. Vortisch, “Modeling intermodal travel behavior in an agent-based travel demand model,” *Procedia Computer Science*, vol. 184, pp. 202–209, 2021, the 12th International Conference on Ambient Systems, Networks and Technologies (ANT) / The 4th International Conference on Emerging Data and Industry 4.0 (EDI40) / Affiliated Workshops.

[4] T. Klinger, “Moving from monomodality to multimodality? changes in mode choice of new residents,” *Transportation Research Part A: Policy and Practice*, vol. 104, pp. 221–237, 2017.

[5] I. Y. (Jackiva), E. B. (Budiloviča), and V. Gromule, “Accessibility to riga public transport services for transit passengers,” *Procedia Engineering*, vol. 187, pp. 82–88, 2017, tRANS-BALTICA 2017: TRANSPORTATION SCIENCE AND TECHNOLOGY: Proceedings of the 10th International Scientific Conference, May 4–5, 2017, Vilnius Gediminas Technical University, Vilnius, Lithuania.

[6] Y.-H. Cheng and S.-Y. Chen, “Perceived accessibility, mobility, and connectivity of public transportation systems,” *Transportation Research Part A: Policy and Practice*, vol. 77, pp. 386–403, 2015.

[7] M. Kawabata and Q. Shen, “Job accessibility as an indicator of auto-oriented urban structure: A comparison of boston and los angeles with tokyo,” *Environment and Planning B: Planning and Design*, vol. 33, no. 1, pp. 115–130, 2006.

[8] I. Benenson, K. Martens, Y. Rofé, and A. Kwartler, “Public transport versus private car gis-based estimation of accessibility applied to the tel aviv metropolitan area,” *The Annals of Regional Science*, vol. 47, pp. 499–515, 2011.

[9] M. A. Saif, M. M. Zefreh, and A. Torok, “Public transport accessibility: A literature review,” *Periodica Polytechnica Transportation Engineering*, vol. 47, no. 1, p. 36–43, 2019. [Online]. Available: https://pp.bme.hu/tr/article/view/12072

[10] M. Givoni and P. Rietveld, “The access journey to the railway station and its role in passengers’ satisfaction with rail travel,” *Transport Policy*, vol. 14, no. 5, pp. 357–365, 2007.

[11] E. J. Molin and H. J. Timmermans, “Context dependent stated choice experiments: The case of train egress mode choice,” *Journal of choice modelling*, vol. 3, no. 3, pp. 39–56, 2010.

[12] K. Bhandari, M. Advani, P. Parida, and S. Gangopadhyay, “Consideration of access and egress trips in carbon footprint estimation of public transport trips: case study of delhi,” *Journal of cleaner production*, vol. 85, pp. 234–240, 2014.

[13] S. Kim, G. F. Ulfarsson, and J. T. Hennessy, “Analysis of light rail rider travel behavior: Impacts of individual, built environment, and crime characteristics on transit access,” *Transportation Research Part A: Policy and Practice*, vol. 41, no. 6, pp. 511–522, 2007.

[14] A. T. Murray, “Strategic analysis of public transport coverage,” *Socio-Economic Planning Sciences*, vol. 35, no. 3, pp. 175–188, 2001.

[15] J. Rychlewski, “Street network design for a sustainable mobility system,” *Transportation research procedia*, vol. 14, pp. 528–537, 2016.

[16] P. A. Singleton, J. C. Totten, J. P. Orrego-Oñate, R. J. Schneider, and K. J. Clifton, “Making strides: state of the practice of pedestrian forecasting in regional travel models,” *Transportation research record*, vol. 2672, no. 35, pp. 58–68, 2018.

[17] R. Moeckel, N. Kuehnel, C. Llorca, A. Moreno, 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.

[18] A. Tønnesen, M. Knapskog, T. P. Uteng, and K. V. Øksenholt, “The integration of active travel and public transport in norwegian policy packages: A study on ‘access, egress and transfer’and their positioning in two multilevel contractual agreements,” *Research in Transportation Business & Management*, vol. 40, p. 100546, 2021.

[19] J. Pucher and R. Buehler, “Walking and cycling for healthy cities,” *Built environment*, vol. 36, no. 4, pp. 391–414, 2010.

[20] D. van Soest, M. R. Tight, and C. D. Rogers, “Exploring the distances people walk to access public transport,” *Transport reviews*, vol. 40, no. 2, pp. 160–182, 2020.

[21] F. Zhen, X. Cao, and J. Tang, “The role of access and egress in passenger overall satisfaction with high speed rail,” *Transportation*, vol. 46, pp. 2137–2150, 2019.

[22] J. Gutiérrez, O. D. Cardozo, and J. C. García-Palomares, “Transit ridership forecasting at station level: an approach based on distance-decay weighted regression,” *Journal of transport geography*, vol. 19, no. 6, pp. 1081–1092, 2011.

[23] R. Cervero, “Walk-and-ride: factors influencing pedestrian access to transit,” *Journal of Public Transportation*, vol. 3, no. 4, pp. 1–23, 2001.

[24] M. Keijer and P. Rietveld, “How do people get to the railway station? the dutch experience,” *Transportation planning and technology*, vol. 23, no. 3, pp. 215–235, 2000.

[25] M. Brons, M. Givoni, and P. Rietveld, “Access to railway stations and its potential in increasing rail use,” *Transportation Research Part A: Policy and Practice*, vol. 43, no. 2, pp. 136–149, 2009.

[26] A. De Witte, J. Hollevoet, F. Dobruszkes, M. Hubert, and C. Macharis, “Linking modal choice to motility: A comprehensive review,” *Transportation Research Part A: Policy and Practice*, vol. 49, pp. 329–341, 2013.

[27] S. Krygsman, M. Dijst, and T. Arentze, "Multimodal public transport: an analysis of travel time elements and the interconnectivity ratio," *Transport Policy*, vol. 11, no. 3, pp. 265–275, 2004.

[28] C.-H. Wen, W.-C. Wang, and C. Fu, "Latent class nested logit model for analyzing high-speed rail access mode choice," *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 2, pp. 545–554, 2012.

[29] M. T. Tran, J. Zhang, and A. Fujiwara, "Can we reduce the access by motorcycles to mass transit systems in future hanoi?" *Procedia-Social and Behavioral Sciences*, vol. 138, pp. 623–631, 2014.

[30] D. R. Loutzenheiser, "Pedestrian access to transit: model of walk trips and their design and urban form determinants around bay area rapid transit stations," *Transportation Research Record*, vol. 1604, no. 1, pp. 40–49, 1997.

[31] Y. Jiang, P. C. Zegras, and S. Mehndiratta, "Walk the line: station context, corridor type and bus rapid transit walk access in jinan, china," *Journal of Transport Geography*, vol. 20, no. 1, pp. 1–14, 2012.

[32] R. Cervero and M. Duncan, "Walking, bicycling, and urban landscapes: evidence from the san francisco bay area," *American journal of public health*, vol. 93, no. 9, pp. 1478–1483, 2003.

[33] P. Zhao and S. Li, "Bicycle-metro integration in a growing city: The determinants of cycling as a transfer mode in metro station areas in beijing," *Transportation research part A: policy and practice*, vol. 99, pp. 46–60, 2017.

[34] A. Polydoropoulou and M. Ben-Akiva, "Combined revealed and stated preference nested logit access and mode choice model for multiple mass transit technologies," *Transportation Research Record*, vol. 1771, no. 1, pp. 38–45, 2001.

[35] T. A. Arentze and E. J. Molin, "Travelers' preferences in multimodal networks: Design and results of a comprehensive series of choice experiments," *Transportation Research Part A: Policy and Practice*, vol. 58, pp. 15–28, 2013.

[36] K. E. Train, *Discrete choice methods with simulation*. Cambridge university press, 2009.

[37] C.-H. Wen and F. S. Koppelman, "The generalized nested logit model," *Transportation Research Part B: Methodological*, vol. 35, no. 7, pp. 627–641, 2001.

[38] C. V. Forinash and F. S. Koppelman, "Application and interpretation of nested logit models of intercity mode choice," *Transportation research record*, no. 1413, 1993.

[39] A. Ermagun and D. Levinson, "Public transit, active travel, and the journey to school: a cross-nested logit analysis," *Transportmetrica A: Transport Science*, vol. 13, no. 1, pp. 24–37, 2017.

[40] Å. Bergman, J. Gliebe, and J. Strathman, "Modeling access mode choice for inter-suburban commuter rail," *Journal of Public Transportation*, vol. 14, no. 4, pp. 23–42, 2011.

[41] G. Debrezion, E. Pels, and P. Rietveld, "Modelling the joint access mode and railway station choice," *Transportation Research Part E: logistics and transportation review*, vol. 45, no. 1, pp. 270–283, 2009.

[42] M. Meng, P. P. Koh, and Y. D. Wong, "Influence of socio-demography and operating streetscape on last-mile mode choice," *Journal of Public Transportation*, vol. 19, no. 2, pp. 38–54, 2016.

[43] D. L. Kurth, C. L. Chang, and P. J. Costinett, "Enhancements to circulator-distributor models for chicago central area based on recently collected survey data," *Transportation Research Record*, vol. 1443, pp. 1443–002, 1994.

[44] R. Schakenbos, L. La Paix, S. Nijenstein, and K. T. Geurs, "Valuation of a transfer in a multimodal public transport trip," *Transport policy*, vol. 46, pp. 72–81, 2016.

[45] L. Puello and K. T. Geurs, "Adaptive stated choice experiment for access and egress mode choice to train stations," in *Proceedings of the World Symposium on Transport and Land Use Research, Delft, the Netherlands*, 2014.

[46] D. A. Hensher and J. M. Rose, "Development of commuter and non-commuter mode choice models for the assessment of new public transport infrastructure projects: A case study," *Transportation Research Part A: Policy and Practice*, vol. 41, no. 5, pp. 428–443, 2007.

[47] M. Rahman, M. S. Akther, and W. Recker, "The first-and-last-mile of public transportation: A study of access and egress travel characteristics of dhaka's suburban commuters," *Journal of Public Transportation*, vol. 24, p. 100025, 2022.

[48] G. Azimi, A. Rahimi, M. Lee, and X. Jin, "Mode choice behavior for access and egress connection to transit services," *International Journal of Transportation Science and Technology*, vol. 10, no. 2, pp. 136–155, 2021.

[49] H. Akaike, "A new look at the statistical model identification," *IEEE transactions on automatic control*, vol. 19, no. 6, pp. 716–723, 1974.

[50] M. Stone, "Comments on model selection criteria of akaike and schwarz," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 276–278, 1979.

[51] L. Creemers, T. Bellemans, D. Janssens, G. Wets, and M. Cools, "Analyzing access, egress, and main transport mode of public transit journeys: evidence from the flemish national household travel survey," 2015.

[52] B. Mo, Y. Shen, and J. Zhao, "Impact of built environment on first-and last-mile travel mode choice," *Transportation research record*, vol. 2672, no. 6, pp. 40–51, 2018.

[53] K. Halldórsdóttir, O. A. Nielsen, and C. G. Prato, "Home-end and activity-end preferences for access to and egress from train stations in the copenhagen region," *International Journal of Sustainable Transportation*, vol. 11, no. 10, pp. 776–786, 2017.

[54] G. Rodrigues and A. Breach, "Measuring up: Comparing public transport in the uk and europe's biggest cities," November 2021, [Online; accessed 20-January-2023]. [Online]. Available: https://www.centreforcities.org/publication/comparing-public-transport-uk-europe-cities/

[55] D. Ariely and J. Levav, "Sequential choice in group settings: Taking the road less traveled and less enjoyed," *Journal of consumer Research*, vol. 27, no. 3, pp. 279–290, 2000.

[56] "Greater manchester travel diary surveys," 2023, [Online; accessed 24-January-2023]. [Online]. Available: https://tfgm.com/trads/

[57] "National travel survey," 2023, [Online; accessed 24-January-2023]. [Online]. Available: https://www.gov.uk/government/collections/national-travel-survey-statistics/

[58] "National travel survey 2019 technical report," 2019, [Online; accessed 25-January-2023]. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/906853/nts-2019-technical-report.pdf/

[59] S. Hess and D. Palma, "Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application," *Journal of choice modelling*, vol. 32, p. 100170, 2019.

[60] "Apollo version 0.1.0," *User manual*, 2019. [Online]. Available: www.ApolloChoiceModelling.com

[61] S. Menard, *Applied logistic regression analysis*. Sage, 2002, no. 106.

[62] D. A. Hensher and W. H. Greene, "Specification and estimation of the nested logit model: alternative normalisations," *Transportation Research Part B: Methodological*, vol. 36, no. 1, pp. 1–17, 2002.

[63] C. Staves, "Physical activity assessment and modelling using household travel surveys," 2020.

[64] R. Moeckel, N. Kuehnel, C. Llorca, A. T. Moreno, and H. Rayaprolu, "Agent-based simulation to improve policy sensitivity of trip-based models," *Journal of Advanced Transportation*, vol. 2020, pp. 1–13, 2020.

[65] 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*, vol. 47, pp. 1787–1808, 2020.

A. Appendix

This appendix contains the model coefficients for the mode choice models for access and egress. To facilitate interpretation, the coefficients are accompanied by a key indicating the level of statistical significance.

Key For Coefficients:
$\cdot$ p-value ≤ 0.1
* p-value ≤ 0.05
** p-value ≤ 0.01
*** p-value ≤ 0.001

Access Mode Choice (McFadden $R^2 = 0.92$) (Bus as main travel mode only)

attribute	value	carD	carP	bus	cycle	walk
INTERCEPT			0.1974***	1.4438***	1.6521***	3.2806***
Urban	False					
	True			0.4199*		0.6007***
Economic status	**1**		-0.1319**		0.2204	0.302**
	2		-0.0763*	0.14.		-0.1515.
	3					
Cars per household			-0.0093*	-0.8196***	-1.3722***	-0.8404***
Age	**1-18**		0.3227***		-0.7939**	0.7058***
	19-29		0.08.	0.5337*		0.706***
	30-49					
	50-59			0.1308	-0.2631.	-0.0308.
	60-69		0.027	-0.4059.	-1.3383**	
	70+				-1.5778**	
Sex	male					
	female		0.021.		-0.3072*	-0.2339*
Driver's license	No					
	Yes		-0.4518***	-1.562***		-0.6043***
Owns bicycle	No					
	Yes			-0.6015***	1.0808***	
Owns season ticket	No					
	Yes		0.1774***	-0.2346.		-0.2007.
Work status	Employed					
	Unemployed				0.4853.	-0.1403
	Retired					-0.292*
Activity type	Mandatory					
	Discretionary				0.2062	0.2224*
	Escort					2.4746***
Log access distance			-0.0013	-0.3365***	-1.6025***	-3.1417***
Trip origin (home)	No					
	Yes		0.0094	-1.1118***	-0.9953***	-1.1056***
NESTING PARAMETERS		0.4297**		0.8307***		

Table A1: Access Mode Choice (Bus is main travel mode)

Access Mode Choice (McFadden R^2 = 0.69) (Train as main travel mode only)

attribute	value	carD	carP	bus	cycle	walk
INTERCEPT			0.1849***	3.3431***	1.2754***	4.5593***
Cars per household			-0.0527***	-0.9129***	-0.8126***	-0.0257
Urban	False					
	True			0.735***		0.2541*
Economic status	1		-0.0807**		0.3659***	-0.1908***
	2		-0.0819***	0.0593		-0.1978.
	3					
Age	1-18		0.4017***		-0.7187***	0.7058***
	19-29		0.169***	0.4024***		0.6686***
	30-49					
	50-59			-0.3401***	-0.7632***	-0.3172***
	60-69		0.049	0.0732.	-0.741***	
	70+				-0.9485*	
Sex	male					
	female		0.1291***	-0.155**	-0.8428***	-0.3717***
Driver's license	No					
	Yes		-0.2667***	-0.644***		-0.5664***
Owns bicycle	No					
	Yes			-0.2713***	2.3401***	
Owns season ticket	No					
	Yes		0.0125	0.0837.		-0.1034*
Work status	Employed					
	Unemployed				0.2002.	0.0478.
	Retired			0.0524.		-0.1046*
Activity type	Mandatory					
	Discretionary				-0.6434***	0.056.
	Escort					0.427*
Log access distance			-0.0645***	-0.7091***	-1.0239***	-2.9425***
Trip origin (home)	No					
	Yes		0.0635*	-1.1448***	-1.4304***	-1.6697***
NESTING PARAMETERS		0.4297**		0.8307***		

Table A2: Access Mode Choice (Train is main travel mode)

Access Mode Choice (McFadden $R^2 = 0.80$)

attribute	value	carD	carP	bus	cycle	walk
INTERCEPT			0.2397***	3.8253 ***	2.6399***	7.5821***
Urban	False					
	True			0.7453***		0.414***
Economic status	1		-0.0993***		0.32**	0.1966***
	2		-0.0739***	0.0761*		
	3					
Cars per household			-0.0442***	-0.9359***	-0.8574***	-0.7267***
Age	1-18		0.3759***		-0.9717***	-0.1735**
	19-29		0.1516***	0.4284***		0.706***
	30-49					
	50-59			-0.2371**	-0.6233***	-0.2321**
	60-69		0.0392*	0.0456 .	-0.8852***	
	70+				-1.1483***	
Sex	male					
	female		0.1051***		-0.6858***	-0.2551***
Driver's license	No					
	Yes		-0.3087***	-0.6694***		-0.5789***
Owns bicycle	No					
	Yes			-0.3055***	1.8442***	
Owns season ticket	No					
	Yes		0.0451**	0.1009*		-0.1374**
Work status	Employed					
	Unemployed				0.3679**	-0.0731 .
	Retired					-0.1325*
Activity type	Mandatory					
	Discretionary				-0.5112***	0.0997*
	Escort					0.9378***
Log access distance			-0.0621***	-0.9627***	-1.5921***	-4.9115***
Trip origin (home)	No					
	Yes		0.0394*	-1.0676***	-1.3009***	-1.58***
Main travel mode	Train					
	Bus			-1.6114***		0.392***
NESTING PARAMETERS		0.1881**		0.5591***		

Table A3: Access Mode Choice

Egress Mode Choice (McFadden $R^2 = 0.91$) (Bus as main travel mode only)

attribute	value	carD	carP	bus	cycle	walk
INTERCEPT			-0.5644***	1.2036 ***	0.7043**	4.3653***
Cars per household			-0.0934.	-1.847***		-0.7261***
Household size			0.1349***	0.1756*		0.1579***
Urban	False					
	True					1.164***
Age	1-18		0.6606***			1.4667***
	19-29		0.0921	1.1506***		0.6336***
	30-49					
	50-59		0.0935	1.0869***		0.4791**
	60-69		0.2847*	0.7775***	-1.1187*	
	70+		0.1283.			
Sex	male					
	female		0.1062	-0.2815	-1.4481***	-0.5467***
Driver's license	No					
	Yes		-0.2561*	-0.35.		-0.6169***
Owns bicycle	No					
	Yes					
Owns season ticket	No					
	Yes			-0.3118.		-0.3038**
Work status	Employed					
	Unemployed		0.5025***	0.8105*		0.3331
	Retired			1.0509***		0.5537***
Activity type	Mandatory					
	Discretionary		0.2586**		-0.55.	-0.0092
	Escort					0.9196*
Log access distance			0.0955.		-0.4307**	-0.8766***
Trip destination (home)	No					
	Yes			-2.0067***		-0.7887***
NESTING PARAMETERS		0.3927**		0.7804***		

Table A4: Egress Mode Choice (Bus is main travel mode)

Egress Mode Choice (McFadden R^2 = 0.59) (Train as main travel mode only)

attribute	value	carD	carP	bus	cycle	walk
INTERCEPT			-0.5267**	-0.0518	-4.4503***	4.1462***
Cars per household			-0.1751***	-1.1774***	-0.8501***	-0.8635***
Household size			0.0553	0.3476***	0.2647***	0.2009***
Urban	False					
	True			1.667***	0.939***	1.4669***
Age	1-18		3.4191***	2.6107***	1.5546*	2.3526***
	19-29		0.9317**	1.1492***		0.867***
	30-49					
	50-59		0.3731**	0.4752***	-0.2544.	
	60-69		0.1669	0.3895**	-1.5958***	
	70+		0.487*	0.7204***	-2.4474*	
Sex	male					
	female		0.4024**	0.1889	-1.1637***	-0.154.
Driver's license	No					
	Yes			-0.7026***		-0.2729***
Owns bicycle	No					
	Yes				4.703***	-0.012
Owns season ticket	No					
	Yes			0.518***		
Work status	Employed					
	Unemployed			-0.3605**		-0.2683**
	Retired			0.1355.		
Activity type	Mandatory					
	Discretionary				-0.4471**	
	Escort		-1.3966.			0.0632
Log access distance			-0.2519**	0.3728*	0.1759	-1.2771***
Trip destination (home)	No					
	Yes		0.1045	-1.9999***	-1.3382***	-2.4505***
NESTING PARAMETERS		0.6178**		0.8426***		

Table A5: Egress Mode Choice (Train is main travel mode)

Egress Mode Choice (McFadden $R^2 = 0.73$)

attribute	value	carD	carP	bus	cycle	walk
INTERCEPT			-0.0925 **	0.9402***	-2.2118***	3.7917***
Household size			0.0403***	0.3053***	0.2785***	0.2124***
Urban	False					
	True			1.6179***	0.8884***	1.4423***
Cars per household			-0.074***	-1.0792***	-0.9134***	-0.8466***
Age	1-18		0.7341***	0.8819***		0.8241***
	19-29		0.3717***	0.8078***		-0.6157***
	30-49					
	50-59		0.1618***	0.2953***	-0.2622*	
	60-69		0.1072*	0.2806***	-1.08***	
	70+		0.162**	0.3865***	-1.86**	
Sex	male					
	female		0.1693***	-0.0843 .	-1.0243***	-0.3176***
Driver's license	No					
	Yes		-0.301***	-0.6385***		-0.4213***
Owns bicycle	No					
	Yes				3.3023***	-0.0414
Owns season ticket	No					
	Yes			0.3338***		
Work status	Employed					
	Unemployed		0.1967**			-0.0395
	Retired			0.2695*		-0.1299 .
Activity type	Mandatory					
	Discretionary			-0.3177***	-0.5045***	-0.1995***
	Escort		-0.9301**			0.1004 .
Log egress distance			-0.1107***	0.0968**	-0.0844	-1.0077***
Trip destination (home)	No					
	Yes		0.1175**	-1.8743***	-1.3284***	-2.1173***
Main travel mode	Train					
	Bus		-0.2269***	-1.1248***	-0.3321***	1.1427***
NESTING PARAMETERS		0.3243**		0.6888***		

Table A6: Egress Mode Choice