Technische Universität München – Professur für Modellierung Räumlicher Mobilität

Prof. Dr.-Ing. Rolf Moeckel

Arcisstraße 21, 80333 München, www.msm.bgu.tum.de

Master's Thesis

Assessing the Impact of Bicycle Highways on Commuter Mode Share in Munich

Author:

Hema Sharanya Rayaprolu

Supervision:

Prof. Dr.-Ing. Rolf Moeckel

Date of Submission: 2017-01-24

Technische Universität München - Professur für Modellierung Räumlicher Mobilität

Prof. Dr.-Ing. Rolf Moeckel

Arcisstraße 21, 80333 München, www.msm.bgu.tum.de

MASTER'S THESIS

of Hema Sharanya Rayaprolu

Date of Issue: 2016-07-25

Date of Submission: 2017-01-24

<u>Topic:</u> Assessing the impact of bicycle highways on commuter mode share in Munich

Major economic centres like Munich attract a significant number of commuters. With the numbers escalating every year, the city battles heavy congestion and overcrowded public transport. One of the city's many attempts to solve traffic issues and promote better quality of life is to foster bicycling in the region. Although the city boasts a bicycle mode share that is above the national average, bicycle trips are mostly short or for leisure. A recent survey conducted to ascertain the preferences and concerns of the city's bicyclists identified two factors that would encourage more and longer bicycle trips – better infrastructure and road safety.

The regional planning authority of the Munich region is testing the feasibility of a bicycle highway network connecting the city and the surrounding districts. Unlike the existing bicycle lanes, bicycle highways allow safer and faster commutes by avoiding intersections with motorized traffic, and offer wider paths that facilitate overtaking. Such infrastructure has been successfully implemented in Denmark and the Netherlands, and is rapidly being embraced in Germany.

With the increasing popularity of electric bikes and growing awareness that the bicycle could solve many traffic related problems, bicycle highways could be the future of long distance commuting. In this context, this thesis aims at building a mode choice model for the region of Munich to gauge the shift in mode share to bicycle. The model will be set up using a synthetic population created from census data and calibrated with the mode shares estimated in the 2008 Mobilität in Deutschland study. Further, an attempt to validate the model using data such as the 2015 Wohnen Arbeiten Mobilität survey data will be made. The model parameters will then be modified to replicate a scenario where the planned bicycle highways are functional and the corresponding mode shares can be calculated. A comparison of mode shares in the base scenario and the scenario with bicycle highways will then be undertaken to assess the impact of bicycle highways.

The student will present intermediate results to the mentor (Prof. Dr.-Ing. Rolf Moeckel) in the fifth, tenth, 15th and 20th week.

The student must hold a 20-minute presentation with a subsequent discussion at the most two months after the submission of the thesis. The presentation will be considered in the final grade in cases where the thesis itself cannot be clearly evaluated.

____

Prof. Dr.-Ing. Rolf Moeckel

Abstract

While traditionally bicycling was restricted to short or leisure trips, people are increasingly bicycling for utilitarian purposes due to growing environmental and health consciousness. Cities are also increasingly investing in promoting bicycling as a measure to mitigate the ill-effects of motorized transportation. Bicycle highways are among such measures aimed to attract more commutes by bicycles. They differ from existing cycling lanes in that they are 3 to 4 m wide, and ensure minimal interaction with motorized traffic by building tunnels or bridges, or by prioritizing bicycle traffic in order to make bicycle commuting faster and safer. With the increasing popularity of pedelecs and e-bikes in Germany, this Dutch concept is being adopted in many German regions.

This thesis aims to assess the potential impact on commuter mode share of bicycle highways proposed for the Munich region. To do this, a discrete choice model was built based on commute data from a national household travel survey conducted in 2008, *Mobilität in Deutschland*. The alternatives investigated in the model include *walk, bicycle, transit* and *auto*. Considering data limitations, the following explanatory variables were selected to specify the model – *age, gender, income, household size, number of children (< 18 years) in household*, *auto availability, bicycle availability* and *travel time*. The model was estimated in a logit modeling framework using R's *mlogit* package. Multinomial logit and nested logit estimations were performed. The estimated model parameters were then applied to scenarios with a pilot bicycle highway proposed in Munich to predict commuter mode shares.

The data appears to lack an adequate level of detail to estimate reliable nested logit models. It was, however, found to be suitable for multinomial logit estimation. The variable *income* was found to be statistically insignificant and hence was removed from the model. The model estimation was found to predict choice probabilities within $\pm$2% of the actual shares when applied to the model dataset. When applied to scenarios with increased average bicycling speeds assumed to result from the bicycle highway, the model's prediction suggests growth in the propensity of bicycle commuting, and a corresponding reduction in the relative attractiveness of auto travel.

The results can be considered to be on the conservative side as the benefits of bicycle highways, besides travel time reductions, could not be included in the prediction. However, the reliability of the results can be improved by validating the estimated model with data from the Munich region, an exercise that could not be performed within the scope of this thesis.

Acknowledgements

I would like to thank Prof. Rolf Moeckel for supervising this thesis work. It would have been impossible to carry out this research within the limited time frame of six months without his continual advice and support. This thesis has benefited a great deal from his immense knowledge in the field of transport demand modeling. I am very grateful to him for his invaluable feedback and constant encouragement.

Many thanks to Dr. Carlos Llorca Garcia and Dr. Ana Tsui Moreno Chou for providing me with all the data I needed during the course of my thesis, and patiently answering to all my questions. Special thanks to Dr. Carlos Llorca Garcia for agreeing to be the second reader of this thesis.

I would also like to thank all my colleagues at the Research Group of Modeling Spatial Mobility for being so compassionate and boosting my morale whenever I was struck somewhere.

Most special thanks go to Josh Tang for his constant support and motivation, and for being there for me at all times. I can’t thank him enough for being so kind as to read the report and help better my writing. I am indebted to him for his endless encouragement and contribution.

Lastly, I would like to express my gratitude to my parents, my little sister and all my friends for all the support and love they have always given me.

Table of Contents

1 Introduction

Bicycling has been experiencing a renaissance over the last few decades. While health and monetary benefits of bicycling have been well known, its societal benefits were recognized only when ill-effects of motorized transportation on the environment surfaced (Banister, 1990). Bicycling is perhaps the most environmentally sustainable of all modes of transport save walking. Its use does not contribute to any form of pollution, and has relatively minimal space, energy and infrastructure requirements (Heinen, van Wee, & Maat, 2010; Kuhnimhof, Chlond, & Huang, 2010). However, it is slower, calls for a greater physical effort, and suffers from a greater exposure to weather and climatic conditions. Nevertheless, growing environmental and health consciousness draws more and more people to bicycle regularly.

With growing levels of congestion and pollution, cities worldwide are increasingly investing in policies to encourage cycling (Heinen et al., 2010). Bicycle highways are one of many such efforts developed with an aim to facilitate bicycling for utilitarian purposes, especially commuting. They differ from other types of cycling infrastructure in that they avoid intersections with motorized traffic, and are wide enough to allow for safe overtaking, thereby increasing cycling speeds (European Cyclists' Federation, 2014). The barrier-free travel eliminates the need for cyclists to stop and accelerate often, facilitating longer distances for the same amount of energy expenditure. The concept was founded in the Netherlands and has spread to many European countries (Tscharnke, 2015).

Recognizing the potential of bicycle highways, many areas in Germany have started studying the feasibility of the implementation of such networks – the Ruhr region is constructing its pilot, routes are being identified in Munich, and many other cities are drafting plans (O'Sullivan, 2016). With the current popularity of bicycling among Germans, great potential for the success of the highways is predicted (Tscharnke, 2015). This thesis aims to assess the probable impact of the bicycle highways proposed in the Munich region on commuter mode share.

Munich, a major economic center in the south of Germany, attracts a significant amount of commuting from its surroundings (Planungsverband Äußerer Wirtschaftsraum München, 2015a). To cope with the increasing demand for transportation, the city incorporates several measures to promote travel by non-motorized and mass transit modes (City of Munich, Department of Urban Planning and Building Regulation, 2006). On the bicycling front, the city aims to be the bicycling capital (Radlhauptstadt, in German) of the country, and runs an extensive campaign to promote regular bicycling (Landeshauptstadt München, n.d. [2016]). Although the city boasts of a bicycle mode share above the national average (14%) (Landeshauptstadt München, Referat für Stadtplanung und Bauordnung, 2010), the trips are mostly short or for leisure (Mühlfenzl, 2016).

A recent survey conducted to ascertain the preferences and concerns of the city's bicyclists identified two factors that would encourage more and longer bicycle trips – better infrastructure and safety (Paul, Bogenberger, & Fink, 2016). In this context, the regional planning authority, Planungsverband Äußerer Wirtschaftsraum, foresees great potential for bicycle highways in the Munich metropolitan region (Planungsverband Äußerer Wirtschaftsraum München, 2015a). This thesis attempts to measure the impact of these bicycle highways on the region's commute mode share. To do this, first, commuter mode choice behavior was simulated based on information from a national household travel survey. This was done by constructing a disaggregate discrete choice model based on a logit modeling framework. The model was then applied to scenarios with bicycle highways proposed for Munich to predict commuter mode shares.

This thesis' research is reported in the rest of the document as outlined below:

- Chapter 2 gives a brief introduction to the concept of bicycle highways. Examples of those existing and being planned in various European regions are discussed.
- Chapter 3 presents the approach adopted in this thesis to conduct the proposed research.
- Chapter 4 details a review of the modeling framework adopted, and previous research concerning similar mode choice studies.
- Chapter 5 describes the various aspects of the commuter mode choice model built – data used (section 5.1), specification of choice set, explanatory variables and model structure (section 5.2), and results of the model estimation (section 5.3).
- Chapter 6 discusses the analysis of scenarios incorporating bicycle highways. Section 6.1 details the dataset used for prediction, and the results of the prediction are illustrated in section 6.2.
- Finally, chapter 7 concludes the report with a retrospective summary of the outcomes, limitations, and recommendations for future research.

2 Bicycle Highways

In the current context, bicycle highways can be defined as ‘fast cycling routes’ which are “high standard bicycle paths reserved for cyclists for fast and direct commuting over long distances” (European Cyclists' Federation, 2014). With the advent of electric bicycles, this type of cycling infrastructure is gaining popularity in many North-European countries as a solution to the mounting problems of congestion and pollution.

The concept was first developed in the Netherlands so cyclists commuting up to 15 km can reach their destinations quickly and safely (Thiemann-Linden & Boeckhout, 2012, p. 1). Copenhagen and London soon followed suit and many other regions including Germany are now catching up with this trend (Tscharnke, 2015).

Owing to the novelty of the concept, there is a lack of universally agreed definition of the characteristics of bicycle highways (European Cyclists' Federation, 2014). Some general characteristics listed by the European Cyclists’ Federation based on individual projects require that bicycle highways –

- Be at least 5km long;
- Have a minimum width – e.g. 3 m, if one-directional, and 4 m, if bi-directional;
- Be separated from motorized traffic and pedestrians;
- Avoid steep climbs and afford mild gradients;
- Avoid frequent stops to enable high average speeds; and
- Provide regular maintenance – e.g. winter service, lighting, service stations, etc.

Although characteristics differ across regions, the criteria governing high quality cycling infrastructure laid out in the national design manuals are consistently followed.

The following sections of this chapter provide an overview of current and planned bicycle highways.

‘Snelfietsroutes’ in the Netherlands

The Netherlands, known to be synonymous with cycling, has had bicycle paths that allow high cycling speeds since the early 1980s (Tscharnke, 2015). The country pioneered the concept of bicycle highways when it inaugurated a 7 km *Snelfietsroute* (fast cycle route) in 2003 (European Cyclists' Federation, n.d.[2014]). *Snelfietsroutes,* also known as *Fietssnelwegen,* are defined as continuous bicycle paths that do not intersect with motorized traffic and hence allow for faster bicycle commutes (Goudappel Coffeng, n.d.a).

These *Snelfietsroutes* were conceived as part of a project launched by the national transport ministry in order to fight congestion (Thiemann-Linden & Boeckhout, 2012, p. 3). The first routes were built alongside motorways in order to entice frustrated auto drivers to shift to bicycles (van der Zee, 2016). These dedicated paths give cyclists total right of way away from main roads and thus allow higher speeds at a relatively low expense of energy (Thiemann-Linden & Boeckhout, 2012, p. 1). They avoid intersection with motorized traffic through underpasses or overpasses wherever possible, and with signal prioritization where the former is not feasible (van der Zee, 2016). A 2010 evaluation of five routes reported a 15% shift to bicycles along the routes due to a reduced travel time (Pfertner, 2015, pp. 10–11).

Figure removed due to possible copyright infringements

Fig. 2.1: Bicycle highways in the Netherlands (Source: Goudappel Coffeng (n.d.b))

The country envisions building a 675 km network by 2025 and currently has more than one-third of the routes in place (European Cyclists' Federation, 2014). Their estimated cost ranges from €0.5 to 2 million per kilometer (including signage and lighting) depending on whether bridges or tunnels are constructed for an intersection-free route (Tscharnke, 2015).

‘Supercykelstier’ in Copenhagen, Denmark

A 500 km network of 28 ‘Supercykelstier’ (super cycle highways) is being realized in the Copenhagen metropolitan area through a cooperation between the capital region and 23 municipalities (Sekretariatet for Supercykelstier, n.d.c, n.d.b). At least 2 of the routes are currently functional and 14 are expected to be completed by 2020. The Danes attribute the success of the project mainly to the cooperation between the municipalities.

The main objective of the project is to create fast, comfortable and safe infrastructure for commuter cyclists, and thereby encourage more people to choose bike over car (Sekretariatet for Supercykelstier, n.d.b). The network is planned to resemble the area’s extensive rail and road network with radial and ring routes sprawling across the city (Sekretariatet for Supercykelstier, n.d.c). The routes are also designed to ensure easy intermodal connections with transit services (Sekretariatet for Supercykelstier, n.d.a). They further afford green waves for cyclists, allowing an average speed of 20km/h even during rush hours. The cost is estimated at €0.11 to 0.36 million per kilometer (European Cyclists' Federation, n.d.[2014]).

Figure removed due to possible copyright infringements

Fig. 2.2: Signal prioritization on Copenhagen’s cycle super highways (Source: Cycle Super Highways, Capital Region Copenhagen (n.d.))

The planners claim that the highways could potentially increase the number of bike commuters in the region by more than 30% (Sekretariatet for Supercykelstier, n.d.b). A 2014 evaluation of one of the routes reported a 52% increase in its use (bicycle counts) since it was opened in 2012 (Sekretariatet for Supercykelstier, 2014). Once complete, the network is expected to save 7,000 tonnes of CO_2 emissions and 40 million Euros in health care costs per year, once functional (Thiemann-Linden & Boeckhout, 2012, p. 2).

'Cycle Superhighways' in London, United Kingdom

London's 'Cycle Superhighways' are one of the city's many efforts to meet its objective of increasing bicycle share by a factor of 5 by 2026 (Thiemann-Linden & Boeckhout, 2012, p. 2). The city carried out extensive market research to develop a system of 12 'Cycle Superhighways' running into the city centre, and approved 11 of them (London Cycling Campaign, n.d.). To date, 6 of the routes have been completed and are in use (Transport for London, n.d.[2016]).

Figure removed due to possible copyright infringements

Fig. 2.3: A stretch of the London's cycle superhighway (Source: Moore (2016))

The project aims to increase commuter cycling by breaking down barriers to commuting by bicycle through a unique package of measures (Transport for London, 2011, p. 1). The highways are designed to run from outer London into and across central London, and offer cyclists safer, faster and more direct journeys into the city (Transport for London, n.d.[2016]). The routes are up to 15 km long and are at least 1.5 m wide in each direction (European Cyclists' Federation, n.d.[2014]). Their estimated total cost is €140 million, which translates to around €1 million per kilometer.

An evaluation of the first two pilot highways launched in July 2010 reports the following results (Transport for London, 2011, p. 2):

- The highways reportedly increased cycling on the routes by 46% and 83% respectively;
- Over 75% of the users were commuters; and
- The journey time was reduced by 5% on average along the two routes.

The project has however been criticized, especially with regard to the name, for being misleading as it implies unimpeded journeys, but in reality, cyclists are often interrupted by intersections and crossing pedestrians (Moore, 2016).

'Radschnellwege' in Germany

German transport and urban planners have been strongly advertising bicycle highways, known as *Radschnellwege,* for years. A study conducted by the German Institute of Urban Local Politicians in 2010 found that *Radschnellwege* will help relieve car traffic along congested routes, making expansion of the road network unnecessary (Tscharnke, 2015).

Germany is unlike the Netherlands where most major cities lie within two hours cycling time of each other (O'Sullivan, 2016). Nonetheless, with more than 2 million Germans owning pedelecs, the idea of bicycle highways has great potential for daily medium length commutes between the inner city and the wider suburban region (Schwägerl, 2016).

Frankfurt is planning a 30 km path south to Darmstadt; Cologne and Hamburg are also in the planning phase; Nuremberg and Berlin are conducting feasibility studies; and Munich is finalizing a 15 km pilot route into its northern suburbs (Bicycling.com, 2015; O'Sullivan, 2016; Schwägerl, 2016). These *Radschnellwege* aim to get commuters out of their cars.

An extension of this idea from paths in metropolitan areas to those connecting cities is being realized in the Ruhr region where a bunch of industrial cities lie scattered at short distances from each other (O'Sullivan, 2016). Another such city-linking bicycle network which also merges with the local cycle networks is planned in the metropolitan region of Hanover-Braunschweig-Göttingen-Wolfsburg (Thiemann-Linden & Boeckhout, 2012, p. 2).

The regions where the highways are being realized (Ruhr region) or are close to realization (Munich region) are discussed in greater detail in the following sub-sections.

The Ruhr region

In November 2015, the Ruhr region opened the first 5 km of the over-100 km long *Radschnellweg*, RS1, set to connect ten cities and four universities (Bicycling.com, 2015; Schwägerl, 2016). When the entire route is complete, expected by 2020, the *Radschnellweg* will bring the cities within a 30 minute cycle distance from each other (European Cyclists' Federation, 2014; O'Sullivan, 2016).

Figure removed due to possible copyright infringements

Fig. 2.4: The planned route of RS1 (Source: Regionalverband Ruhr (n.d.))

Unlike the existing single-lane bike paths, the *Radschnellwege* are 4 to 5 m wide, provide overtaking lanes in both directions, have high-quality asphalt pavements, and avoid traffic signals and road crossings with the help of overpasses and underpasses (Bicycling.com, 2015; Schwägerl, 2016). They are fully lit, have a parallel pedestrian path and will be cleared of snow in winters (van der Zee, 2016).

The RS1 is intended to serve the almost 2 million people that live within 2 km of the route for daily commuting (Bicycling.com, 2015). It has been planned to parallel the A40 in order to entice the region's commuters to forsake cars and ride bicycles (Schwägerl, 2016). In such a densely populated polycentric region, the highways are anticipated to impose a significant relief on road and rail congestion (O'Sullivan, 2016). The regional association predicts that the 100 km path could take 52,000 cars off the roads every day, reducing daily car kilometres by around 400,000 km and annual CO_2 emissions by 16,600 tonnes (Ruhr Regional Association, n.d.[2016], p. 4).

A feasibility study conducted by the Ruhr Regional Association (n.d.[2016], pp. 20–23) estimates the cost of RS1 at €183.7 million, corresponding to a €1.81 million per kilometre cost. It also estimates a benefit-cost ratio of 4.8 for their establishment by considering construction and maintenance costs, and benefits in monetary units of emission reductions, improved safety and contribution to public health.

The funding of the first 5 km stretch was shared by the European Union (50%), the state of North Rhine-Westphalia (30%) and the regional authority (20%) and the sources for the rest of the path are not yet in place (Bicycling.com, 2015; O'Sullivan, 2016).

Munich metropolitan region

The establishment and expansion of bicycle networks in the Munich region have been traditionally based on the belief that cycling is purely a leisure activity (Mühlfenzl, 2016). However, with the increasing knowledge of pros of bicycling, and the growing popularity and affordability of pedelecs and e-bikes, more and more people are considering cycling to work. But the area lacks a network suitable for commuters. To address this, the planning association of the Greater Munich region, PV (Planungsverband Äußerer Wirtschaftsraum), identified enormous potential for 14 *Radschnellwege* in the region that can enable speeds up to 30 km/h.

The 14 protected two-lane highways recommended by PV extend from the city of Munich into the surrounding suburbs within a radius of 20 km (Mühlfenzl, 2016; O'Sullivan, 2016). They are suggested to be 4 m wide and avoid intersections with motorized traffic wherever possible (Walter, 2015). No helmet or minimum speed regulations are intended. The routes are faster only because intersections with traffic would be avoided.

Figure removed due to possible copyright infringements

Fig. 2.5: Recommended and shortlisted bicycle highways in Munich region (Source: Planungsverband Äußerer Wirtschaftsraum München (2015a, p. 51))

Experts explored the 14 corridors in a study commissioned by the city of Munich and four other counties (Walter, 2015). They shortlisted six of them (yellow routes in Fig. 2.5) and recommended a 17 km pilot route which would run between the city boundary and adjacent municipalities to the North (Schwägerl, 2016). However, the process has been delayed due to the complexity of the project and the unclear division of responsibilities in the state (Mühlfenzl, 2016).

In Germany, while the federal government builds and maintains roads, railways and waterways, local authorities are responsible for cycling infrastructure (Bicycling.com, 2015). The limited financial sources available to local governments impede the execution of projects of the scale of bicycle highways. In the Ruhr region, the state (North Rhine-Westphalia) assigned bicycle highways a status equivalent to rural roads, and assumed the responsibility of their construction and maintenance (Mühlfenzl, 2016). Planners assert that without the active participation of the state, land acquirement issues could further delay the realization of the bicycle highways.

The German bicycle club, ADFC, expects costs up to €1 million per km and suggests part contribution from the regular road budget as the bicycle highways also relieve the roads (Walter, 2015).

Other lesser known bicycle highways

'Fietsostrade' in Belgium

Belgium's bicycle highways known as 'Fietsostrade' are almost exclusively limited to the Dutch speaking region, Flanders (European Cyclists' Federation, n.d.[2014]). An extensive, well developed network of bicycle highways is being developed in the Provinces of Antwerp, Flemish Brabant, East Flanders and West Flanders, among others. The project is being funded by the regional state of Flanders (40%), the provinces (40%) and municipalities (20%). As the individual provinces are responsible for the development of the infrastructure, collective information on the length of routes planned, progress, etc. are currently unavailable.

'Vélostras' in Strasbourg, France

Strasbourg is awaiting the completion of a 130 km network of fast cycling routes called 'Vélostras' by 2020 (European Cyclists' Federation, n.d.[2014]). A 12 km route has been completed. The network is planned to constitute 3 ring routes and 9 radial routes connecting the centre of the city and its surroundings. The local sections of the European cycle route network, Eurovelo, are also planned to be integrated into the network. The infrastructure is being developed to facilitate an average speed of 20km/h. The city desires to increase its bicycle mode share to 16% by 2025 with the help of the Vélostras.

'Supercykelväg' in Sweden

A 16 km fast cycling route called 'Supercykelväg' is being constructed to connect the Swedish cities of Malmö and Lund (European Cyclists' Federation, n.d.[2014]). The route is being developed with four lanes to allow for safe overtaking in both directions. As Scandinavian winds can be a hindrance for people to take up long distance cycling, wind protection is additionally being ensured on the highway. A €0.34 million cost per kilometer is estimated.

‘Super-Sykkelveier’ In Norway

The Norwegian government recently announced its plan to build bicycle highways in and around nine of its cities and is set to invest £700 million in their construction (van der Zee, 2016). The *Super*-Sykkelverier are envisaged as an important means to increase bicycling and combat the country’s struggle with emissions.

‘Velobahnen’ in Switzerland

The Swiss bicycle highways, *Velobahnen*, are considered analogous to Autobahns. The city of St. Gallen is considering a *Velobahn* running east-west across the city (SRF Schweizer Radio und Fernsehen, 2015). The city of Winterthur is also planning five such dedicated and separated routes targeted towards commuting cyclists and expects completion of the first in 2018 (20minuten, 2014).

3 Approach

The focus of this thesis is the bicycle highway network that is being planned for the Munich region. The idea is to use a mode choice modeling framework to estimate the impact of the bicycle highways on the region’s commuter mode share. This approach involves two exercises:

1. Formulation of a suitable mode choice model; and
2. Prediction of choice probabilities in a scenario with bicycle highways using the mode choice model.

To carry out these exercises, the following steps have been adopted:

- Review of the choice modeling framework;
- Review of previous studies on commuter bicycle choice models;
- Exploration of available data;
- Specification of model choice set, variables, and structure;
- Estimation of model parameters and analysis of results;
- Dataset preparation for prediction; and
- Scenario analysis for the prediction of impacts of bicycle highways using the model.

A detailed description of the above steps is given in the following chapters.

4 Review of Mode Choice Modeling Research

Modeling mode choice is an essential element of the travel demand modeling framework. For trips between two locations, mode choice models predict the share of trips between the different modes in consideration. In the current state of affairs, with increasing restrictions on energy and space consumption, measures to encourage travel by mass transit and non-motorized modes are rising. Mode choice models can be useful tools in predicting the impacts of such measures.

Mode choice analyses gained momentum in the 1960s and 70s when a shift occurred from unimodal analyses for capital investment decisions to multimodal planning and policy considerations for mass transit operations (Ben-Akiva & Lerman, 1985, p. 1; Ortúzar & Willumsen, 2011, p. 227). Until the early 1980s, most studies employed simple regression techniques, and analyzed the travel behavior of groups of individuals aggregated based on geographical zones. These aggregate models, however, lacked precision in emulating reality due to their inability to handle a range of causal factors, and often suffered biases due to correlation between aggregated elements (Muñoz, Monzon, & Daziano, 2016; Ortúzar & Willumsen, 2011, p. 229).

The 1980s saw an increasing uptake of disaggregate modeling (Ortúzar & Willumsen, 2011, p. 227). While aggregate models observe groups of travelers, disaggregate models analyze travel choices of each individual, or individual household, and therefore can better capture choice relations through a wider range of explanatory variables. As most travel surveys collect data at either the individual or the household level, most analysts prefer disaggregate modeling (Ben-Akiva & Lerman, 1985, p. 2).

Disaggregate models are based primarily on two types of data – revealed preference (RP) and stated preference (SP) (Hensher, Rose, & Greene, 2005, pp. 5–6). RP data records actual choices made by individuals in real market situations, and SP data reports choices that individuals would make in hypothetical choice situations. While RP data is a real world representation and hence more reliable, SP data is especially useful in predicting choice among existing and new alternatives. Studies based on SP data, however, involve the execution of carefully designed surveys in order to ensure reliable responses (Parkin, Wardman, & Page, 2008, p. 96).

At the disaggregate level, individual preferences have been observed to be best described in a discrete choice framework where individuals' choices are explored using the principle of utility maximization (Ben-Akiva & Lerman, 1985, p. 2). A brief overview of this framework is presented in the subsequent section of this chapter, followed by a review of past research in this area. Given the broad nature of this topic, only relevant studies of commuter mode choice that explore bicycle choice were considered.

4.1 Discrete Choice Modeling Framework

Within the discrete choice framework, individuals' choice behavior is modeled econometrically using the principle of utility maximization (Ben-Akiva & Lerman, 1985). In other words, individuals are modeled to choose the alternative with the highest utility when confronted with a set of alternatives. Here, utility is the index of attractiveness of an alternative, and is expressed as a function of the alternative's attributes. This theory is based on the assumptions that individuals behave rationally, possess complete information about all alternatives, and are faced with a mutually exclusive and collectively exhaustive choice set.

A discrete choice model is executed in two broad steps. First, each alternative is assigned a utility in the form of a parameterized function described by its attributes (observable independent variables) and unknown parameters (Ben-Akiva & Lerman, 1985). These unknown parameters are then estimated from a sample of observed choices made by a set of individuals in similar choice situations. These two steps – model specification and model estimation –constitute the modeling framework.

While specifying a model, an important challenge is to decide the functional form of the utility function (Ben-Akiva & Lerman, 1985). For computational convenience, most analysts assume an additive, linear-in-parameter function. The theory requires modelers to assume one universal choice set and to assign each individual decision maker an individual choice set (a subset of the universal choice set) based on her or his individual income and time budgets (Ben-Akiva & Lerman, 1985, p. 48). Generally, these individual constraints are addressed by introducing a vector of socio-economic characteristics into the utility function, and common choice sets are assigned across homogenous samples.

Under this framework, all individuals with the same attribute values and similar socio-economic characteristics would always choose the same alternative – the one with the highest utility (Ortúzar & Willumsen, 2011, p. 230). However, in reality, we see differences in the choices of similar individuals because individuals are not perfectly rational, and models lack perfect information. To address the irrationalities, a probabilistic choice approach based on random utility theory is adopted by adding some randomness to the utility. Consequently, the utility function of an alternative i for an individual n, U_{in}, is expressed as a combination of a measureable, systematic component, V_{in}, and a random component, ε_{in}, as shown in Equation 4.1.

Equation 4.1

$$U_{in} = V_{in} + \varepsilon_{in} \qquad \text{(Ortúzar \& Willumsen, 2011, p. 230)}$$

In this equation, the systematic component, V_{in}, is the parameterized function of the observable attributes of the alternative i which can be written as in Equation 4.2.

Equation 4.2

$$V_{in} = \beta_{1i}X_{1in} + \beta_{2i}X_{2in} + \cdots + \beta_{ki}X_{kin}$$ (Ortúzar & Willumsen, 2011, p. 231)

Where, $X_{1in}, X_{2in}, \dots, X_{kin}$ are the k independent variables that include both attributes of the alternative i and socio-economic variables of the individual n; and

$\beta_{1i}, \beta_{2i}, \dots, \beta_{ki}$ are the unknown parameters which are assumed to be constant across individuals, and may vary across alternatives.

In this utility, the random component, ε_{in}, caters to the unobserved taste variations among individuals and other observational errors (Ortúzar & Willumsen, 2011, p. 231). It is assumed as a random variable following a certain probability distribution function. The utilities of the alternatives are therefore random variables as well. Therefore, the probability that an alternative is chosen is now the probability that it has the greatest utility among all the available alternatives (Ben-Akiva & Lerman, 1985, p. 3). Considering a choice set containing two alternatives i and j, the probability that the individual n chooses the alternative i, P_{in}, is the probability that the utility of alternative i, U_{in} is greater than or equal to the utility of alternative j, U_{jn}.

Equation 4.3

$$P_{in} = Probability\,\{\, U_{in} \geq \, U_{jn} \,\}$$ (Ben-Akiva & Lerman, 1985, p. 65)

When the utilities in the above equation are expressed in terms of their systematic and random components, probability P_{in} can be written as in Equation 4.4.

Equation 4.4

$$P_{in} = Probability\,\{(\varepsilon_{in} - \varepsilon_{jn}) \, \geq (V_{jn} - V_{in})\}$$ (Ben-Akiva & Lerman, 1985, p. 65)

Where, V_{in} is the systematic component of utility U_{in};
ε_{in} is the random component of utility U_{in};
V_{jn} is the systematic component of utility U_{jn}; and
ε_{jn} is the random component of utility U_{jn}.

In order to compute probabilities as described above, a certain probability distribution is assumed for the random error components (ε). The β parameters are then estimated using a maximum likelihood estimation (Ben-Akiva & Lerman, 1985).

Different assumptions about the distribution of ε have led to different types of discrete choice models (Ben-Akiva & Lerman, 1985). Among the various types of models, the logit class of models which assume a logistic distribution of the error components are most widely used. Within the logit class, multinomial logit and nested logit models are the most popular discrete choice models employed in travel demand forecasting (Munizaga & Ortúzar, 1999), and are briefly reviewed here.

4.1.1 Multinomial Logit

Multinomial logit (MNL) models are logit models applied to choice sets containing more than two alternatives. Like all logit class models, they assume a logistic distribution for the set of error components. This is equivalent to assuming that all individual error components ($\varepsilon_{in}, \varepsilon_{jn}, \ldots$) are independent and identically Gumbel distributed with a location parameter η and scale parameter $\mu > 0$ (Ben-Akiva & Lerman, 1985, p. 71). By adding a constant term to the systematic utility components of the alternatives, the location parameter η is set to null. Under these assumptions, the probability that an individual n chooses an alternative i from a choice set C_n, P_{in}, is expressed by the MNL model as in Equation 4.5.

Equation 4.5

$$P_{in} = \frac{\exp(\mu V_{in})}{\sum_{j \epsilon C_n} \exp(\mu V_{jn})}$$ (Ben-Akiva & Lerman, 1985, p. 103)

Here, the systematic utilities of alternatives, V_{in} and V_{jn}, are linear-in-parameter as in Equation 4.2 with an additional constant term, and for convenience, the scale parameter μ is set to 1.

An important and widely discussed aspect of the MNL model is the IIA property (Independence from Irrelevant Alternatives property) which states that when choice probabilities of two alternatives are non-zero, their ratio is independent of any other alternative in the choice set (Ortúzar & Willumsen, 2011, p. 234). This property is evident from Equation 4.5. By virtue of this property, the addition or removal of an alternative to or from the choice set has the same effect on every other alternative in the choice set. But in reality, alternatives are not completely independent, for example, when new transit operations are introduced in a city's transport offer, current captive riders are more likely to switch to the new service than current captive drivers are. This inability of MNL models to capture correlations between alternatives is addressed in the nested logit modeling framework. Nevertheless, MNL models are still applied in the majority of mode choice models owing to their simple mathematical form, and ease of estimation and interpretation (Koppelman & Bhat, 2006, p. 157).

4.1.2 Nested Logit

Nested logit (NL) models were designed, in part, to solve the problem of having correlated alternatives in MNL models (Munizaga & Ortúzar, 1999, p. 26). In NL models, correlated alternatives are grouped into hierarchically arranged nests so that all alternatives within a nest are similarly correlated. For instance, auto and transit modes could be grouped into one nest called 'motorized', and walking and bicycling could be under a 'non-motorized' nest.

The formulation of the model clarified by Munizaga and Ortúzar (1999) is described here. Consider a model with two hierarchical levels. Say, i represents a nest, or, an alternative at the upper level, and j represents an alternative within nest i, i.e., an alternative of the lower level. The utility function can be expressed as the sum of a component representative of the nest i, U_i, and a component representing the effect of alternative j within nest i, $U_{j/i}$.

Equation 4.6

$$U(i,j) = U_i + U_{j/i}$$ (Munizaga & Ortúzar, 1999, p. 26)

Decomposing the above utility into systematic and random components, the above equation can be expressed as –

Equation 4.7

$$U(i,j) \;=\; V(i,j) \;+\; \varepsilon(i,j)$$ (Munizaga & Ortúzar, 1999, p. 26)

The random component of the above equation can also be decomposed into a component common to all alternatives in nest i and a component specific to each alternative j as follows –

Equation 4.8

$$\varepsilon(i,j) = \varepsilon_i + \varepsilon_{j/i}$$ (Munizaga & Ortúzar, 1999, p. 26)

Say, the random component pertaining to the upper level with nests has a scale factor β and that pertaining to the alternatives within nest i has λ_i as the scale factor (Fig. 4.1). Then, the probability of choosing nest i, P_i, and the probability of choosing alternative j which belongs to nest i, $P_{j/i}$, are expressed as in Equation 4.9 and Equation 4.10 respectively.

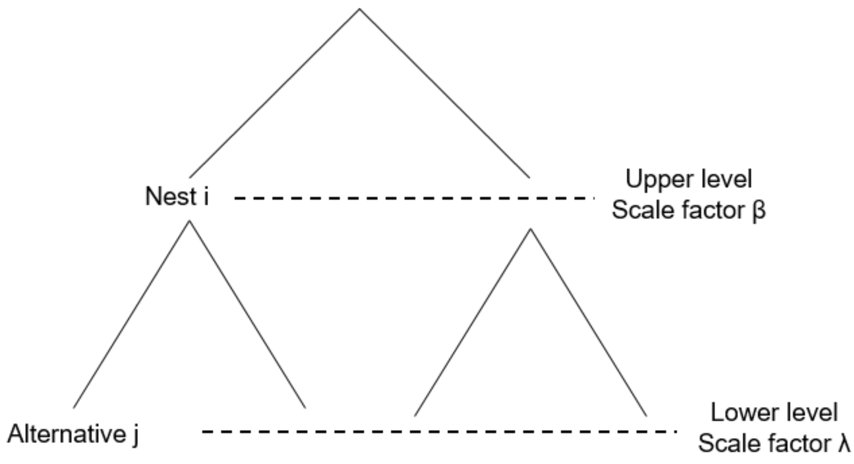

Fig. 4.1: A nested model with two hierarchical levels

Equation 4.9

$$P_i = \frac{\exp(\beta V_i)}{\Sigma_{I \epsilon R} \exp(\beta V_I)}$$ (Munizaga & Ortúzar, 1999, p. 26)

Where, I belongs to the set of a real numbers R, and indicates the number of nests, or number of alternatives in the upper level.

Equation 4.10

$$P_{j/i} = \frac{\exp(\lambda_i V_{j/i})}{\Sigma_{k \epsilon C(i)} \exp(\lambda_i V_{k/i})}$$ (Munizaga & Ortúzar, 1999, p. 26)

Where, k represents all alternatives within nest i.

The probability of choosing nest i, and within it, alternative j, P_{ij}, is then obtained by the product of P_i and $P_{j/i}$ (Munizaga & Ortúzar, 1999, p. 26).

Under this framework, the systematic utilities of the alternatives of the lower level, $V_{j/i}$, are established through linear-in-parameter equations similar to the MNL model. The utilities of the nests at the upper level, V_i, are obtained by a logsum as in the following Equation 4.11.

Equation 4.11

$$V_i = \emptyset_i \, ln \, \Sigma_{k \epsilon C(i)} \exp(\frac{V_k}{\emptyset_i})$$ (Munizaga & Ortúzar, 1999, p. 27)

Where, k represents the alternatives within nest i ,and $\emptyset_i = \frac{\beta}{\lambda_i}$ is the structural parameter corresponding to each nest (Munizaga & Ortúzar, 1999, p. 27), and is called the nesting coefficient, or logsum parameter. For computational convenience, β is generally set to unity, setting the nesting coefficient, $\emptyset_i$, to $\frac{1}{\lambda_i}$. For correlated alternatives within a nest, the value of $\emptyset$ lies between 0 and 1 (Koppelman & Bhat, 2006, p. 163). A perfect correlation between pairs of alternatives occurs when $\emptyset = 1$. The smaller the value of $\emptyset$, the greater the substitution among alternatives in the nest.

By allowing for correlation between alternatives, the NL framework generates models with more reliable choice elasticities than MNL models. Nevertheless, the added complexity places greater computational requirements.

4.2 Commuter Bicycle Choice Models

The choice of mode to travel to work has been extensively researched. However, the majority of the commuter mode choice studies (Hensher & Bullock, 1979; Hensher & Rose, 2007; Hess, 2001; Schwanen & Mokhtarian, 2005; Washbrook, Haider, & Jaccard, 2006) do not extend their choice set to include bicycling as an independent alternative. They most commonly measure the impacts of policies affecting travel times and costs by auto and transit. This could be attributed to the lack of such research in Europe (Kuhnimhof et al., 2010) and the near-negligible bicycling mode share outside of Europe (Twaddle, 2011). Although a lot of research has gone into determining factors influencing bicycle choice (Bergström & Magnusson, 2003; Dunlap, 2015; Fernández-Heredia, Jara-Díaz, & Monzón, 2016; Heinen, van Wee, & Maat, 2010; Maldonado-Hinarejos, Sivakumar, & Polak, 2014; Muñoz, Monzon, & Daziano, 2016; Nankervis, 1999; Sener, Eluru, & Bhat, 2009; Tilahun, Levinson, & Krizek, 2007; Willis, Manaugh, & El-Geneidy, 2015), few studies explore bicycling for utilitarian purposes. Such studies investigating bicycle choice for commuter trips are reviewed here.

Noland and Kunreuther (1995) conducted a RP study of commutes of both bicyclists and the general public in Philadelphia to deliver policy recommendations to promote bicycling. They built a MNL model to understand how individuals' mode choice is affected by their perceptions about travel costs, convenience, comfort and risks of bicycling. Their investigation concluded that two courses of action would be needed to increase bicycle commuting in the region – providing suitable cycling infrastructure for results in the short-run; and implementing policies that denigrate the relative convenience of automobile commuting for sustainable results in the long run.

In another study in the UK, Wardman, Hatfield, and Page (1997) investigated if the country's cycling strategy would be successful in meeting its target through a SP survey of 221 employed individuals. They designed scenarios with varying types and levels of cycling infrastructure, and assessed the results through a MNL model. They found that unsegregated cycling infrastructure would not attract many cycling trips, and even segregated cycling infrastructure would only help achieve the target of quadrupling cycle share in the most favorable set of circumstances. With these results, they stressed the need for traffic management, and restraint measures in addition to segregated cycling infrastructure to achieve the country's goals.

Rodríguez and Joo (2004) studied commuting preferences of students and staff at the University of North Carolina to identify the impact of local physical environment on their choice to walk and bike. They measured local physical environment in terms of topography, sidewalk availability, residential density and presence of walking and cycling paths. By employing an NL model, they found local topography and sidewalk availability to be highly significant for the choice of non-motorized modes.

Through a large, comprehensive study, Wardman, Tight, and Page (2007) investigated factors influencing the propensity to cycle to work in the UK. They observed that most RP studies are limited to examining existing facilities, and SP studies often engender response biases and might seldom be suitable for forecasting owing to inappropriate scales. By methodically combining both RP and SP data from different sources, they built a logit model to predict the impacts of different measures to encourage cycling. They tried different hierarchical logit models and did not find a valid reason to deviate from a MNL model. Their analysis explored the potential of various types of cycling infrastructure, employer payments for cycling to work, work-site facilities, perceptions and attitudes, and other socio-economic variables, and found that a completely segregated cycle-way induces a 55% increase in commuters cycling and 3% decrease in car use to work as compared to the base situation.

Dill and Wardell (2007), Buehler (2012), Hamre and Buehler (2014) and Heinen, Maat, and van Wee (2013) dedicated their research to evaluating the effect of work-site amenities like provision of showers, bicycle parking, employer benefits or payments for bicyclists etc. on encouraging commuting by bicycling. Heinen, Maat, and van Wee (2011) studied the role of attitudes on the choice to commute by bicycle over various distances.

In contrast to most recent travel behavior studies which undertake disaggregate analyses, Parkin et al. (2008) conducted an aggregated study to estimate the determinants of bicycle mode share for commutes. They contest that although disaggregate studies are powerful in studying choice behavior in detail, they overestimate the impacts, especially when based on SP data, due to possible response biases. They explored an aggregate model to complement an existing disaggregate one as certain attributes such as availability of infrastructure, slope grade, hilliness, density etc. are better measured on an aggregate level. However, they admit that attributes explaining traveler's psychological factors are better addressed by disaggregate models. With a binary logit modeling approach, their prediction of the impacts of a completely segregated cycle-way are rather modest in comparison to the prediction of Wardman et al. (2007) that segregated cycle-ways would result in a 55% increase in bicycle trips. Their finding supports their assertions about the downsides of using SP data to forecast impacts.

While most bicycle choice studies employed the modeling framework to predict impacts of policy measures, some studies (Bowman, 2014; Broach, 2016; Hunt & Abraham, 2007) attempted to improve modeling efficiency by integrating a route choice module into the mode choice modeling framework. As the majority of investments for bicycle transportation go into infrastructure improvements, a route choice model would better capture impacts of such measures.

Bicycle commuting has also drawn interest from fields beyond travel demand forecasting. Engbers and Hendriksen (2010) conducted a health impact study on commuter cyclists in the Netherlands, and Buekers, Dons, Elen, and Panis (2015) studied the health impacts of commuter mode shift from cars to bicycles in Belgium by applying their model to two Belgian bicycle highways. Sigurdardottir, Kaplan, Møller, and Teasdale (2013) carried out a behavioral study of Danish adolescents to understand their intentions to commute by car or bicycle as adults.

In conclusion, predicting the impact of infrastructural improvements on commuter mode share has been well researched. However, there is a lack of consensus on many aspects including the level of aggregation, the type of data to be used, and the modeling framework. Furthermore, studies investigating the impacts of improved cycling infrastructure most commonly forecast hypothetical scenarios. To my knowledge, there are no studies (documented in English) that predict the impacts of planned cycling infrastructure like bicycle highways on commuter mode shares. In this context, this thesis aims to conduct a disaggregate study based on RP data through a logit modeling framework to estimate the change in mode share induced by bicycle highways. Based on the findings, some of the modeling issues would be revisited.

5 The Modeling Framework

The current model adopts a discrete choice modelling framework described in section 4.1, and is based on data from a German household travel survey. The construction of the model and its functioning is explained in this chapter. Section 5.1 gives an overview of the data used, and sections 5.2 and 5.3 demonstrate the model specification and estimation respectively.

5.1 Dataset

The model's dataset is extracted from the German household travel survey, *Mobilität in Deutschland*, conducted in 2008. In this section, first an overview of the survey and its contents are presented, followed by a briefing on the criteria used to select the final dataset to specify the model.

5.1.1 Mobilität in Deutschland 2008

Mobilität in Deutschland (MiD), which translates to 'Mobility in Germany', is a nationwide travel survey organized by the Federal Ministry of Transport, Building and Urban Development (Bundesministerium für Verkehr, Bau- und Stadtentwicklung) (Follmer et al., 2010). It is a repeated, cross-sectional survey conducted sporadically (once every 6 to 10 years) over an entire year in order to observe daily travel behavior trends of individuals and households. The survey has been carried out five times since 1976 – KONTIV[1] 1976, KONTIV 1982, KONTIV 1989, MiD 2002 and MiD 2008. The latest survey, MiD 2016, is currently under progress (infas Institut für angewandte, n.d.).

MiD 2008, the most recent completed survey, records mobility-related and socio-demographic information of over 25,000 households with around 100,000 household members of all age groups (German Aerospace Center, 2012). Households are meticulously, yet randomly, selected to ensure an accurate representation of the country's demographics through adequately calculated weights and expansion factors. The survey was carried out in two phases – a household survey collecting particulars of the household, household members, means of transport available, and other features reported by one household member, followed by personal interviews of all household members about their general mobility and a travel journal of their trips on a fixed date (Follmer et al., 2010). A broad overview of the information collected in both phases of MiD 2008 is depicted in Fig. 5.1.

[1] The survey was carried out under the name KONTIV ('Kontinuierliche Erhebung zum Verkehrsverhalten' which translates to 'Continuous Survey on Travel Behavior') in the former West Germany. It was renamed as MiD post-unification.

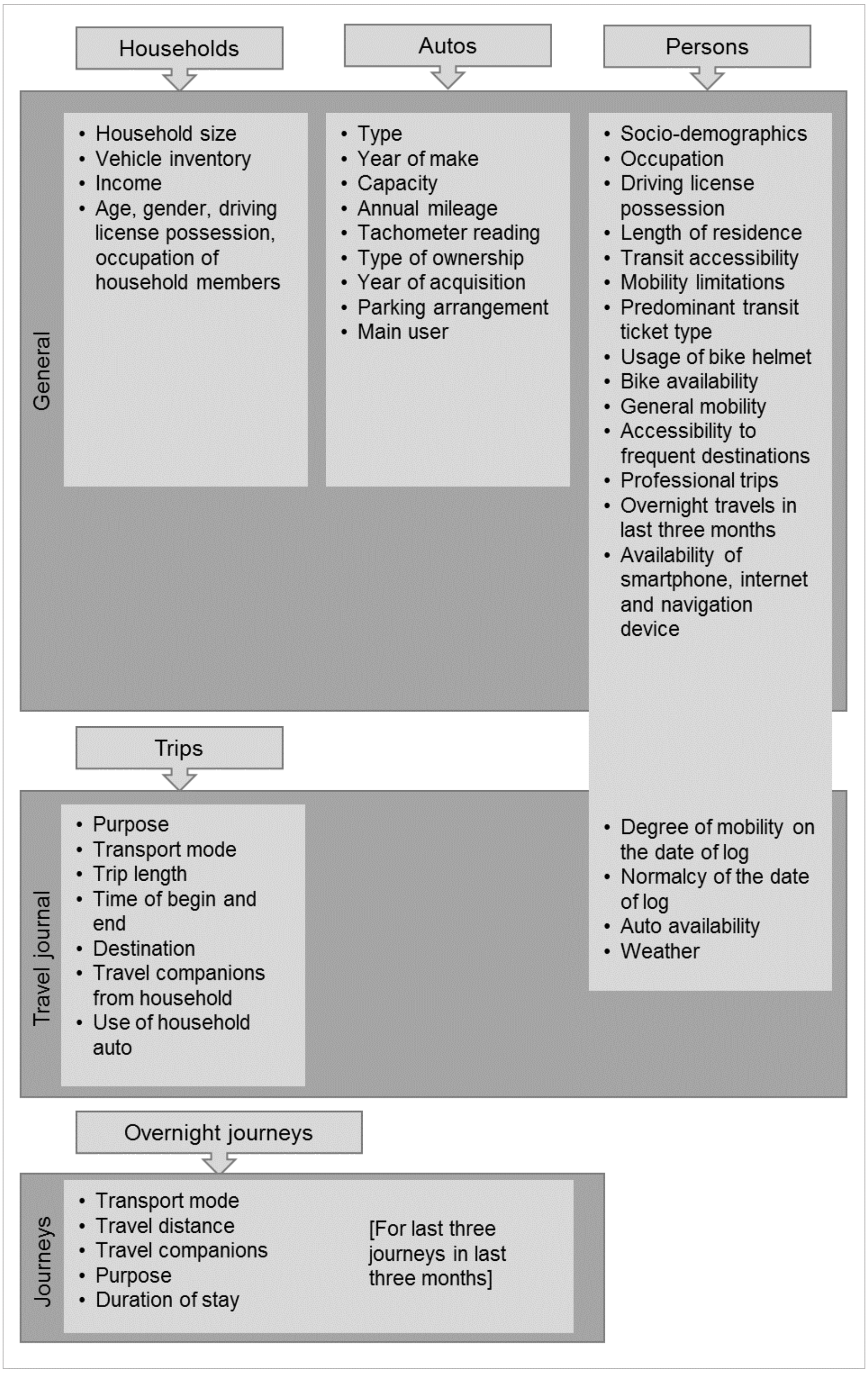

Fig. 5.1: Overview of MiD 2008 data (translated from MiD 2008 documentation)

Owing to its large sample size, generalizability, availability, and lack of a more recent data source, MiD 2008 has been used in this study assuming minimal change in household travel behavior. The data and documentation of MiD2008 was availed from the 'Modeling Spatial Mobility' research group for the purpose of this study. For reference, the documentation (in German) can be found at Bundesministerium für Verkehr und digitale Infrastruktur (n.d.[2009]).

5.1.2 Model Dataset

MiD 2008 data records information on various travel-related aspects from all over the country. As the current model is intended for commute trips in the Munich region, a quick analysis of the survey records was exercised to identify a dataset most relevant to the model's purpose. The records were filtered based on two main criteria – trip purpose, and locational properties.

As a first step, trips from home to work were extracted from the complete set of records. Individuals from 8,357 households out of all the 25,922 surveyed households reported trips from home to work in their travel journals.

The natural next step would be to extract the households located in the Munich region from these 8,357 households. But, MiD 2008 data does not offer such a level of resolution. Household locational characteristics are limited to the Federal states and region types based on population density. To overcome this limitation, households located in the urban regions of Bavaria were considered: 616 home-to-work trips were found to be reported by individuals from 454 households in the urban regions of Bavaria. Considering that the number of trips (616) is not large enough to base a functional model on, trips from urban areas all over Germany were selected instead. This resulted in a final dataset of 8,192 home-to-work trips of individuals from 6,215 households from all urban regions in Germany, with each trip corresponding to one individual. This dataset of 8,192 commute trips was used to specify the model, along with the corresponding sample weights recommended by the survey.

5.2 Model Specification

Within the discrete choice modelling framework, specifying a model involves identifying the choice set, selecting explanatory variables, and deciding the model structure (MNL, NL, etc.) (Ortúzar & Willumsen, 2011, p. 269). To accomplish these steps, this section draws greatly from the mode choice modeling research reviewed in chapter 4.

5.2.1 Choice Set

The choice set of the model was selected based on the travel modes recorded in MiD 2008 data while keeping the purpose of the model in mind.

The MiD 2008 trip data recorded all modes used by individuals during each trip, and categorized them into six main transport modes, among others. While reporting trips in travel journals, individuals selected the modes they used from a comprehensive list of 18 modes. In their analysis of the reported trips, MiD 2008 identified one main mode for each trip, grouped the modes into broader categories, and further condensed them into six modal categories. For trips with multiple modes, the main mode was assigned based on a hierarchy of modes developed based on travel distances. The 18 modes of choice arranged in the hierarchical order, an intermediate categorization into 12 groups, and the final categorization into 6 modes is illustrated in Table 5.1.

Table 5.1: Categorization of modes in MiD 2008 (adapted and translated from documentation)

Mode hierarchy	**Main mode category (12)**	**Main mode category (5)**
1 – Airplane	1 - Ship | train | coach | airplane	1 – Long distance transit
2 – Bus coach	1 - Ship | train | coach | airplane	1 – Long distance transit
3 – Train	1 - Ship | train | coach | airplane	1 – Long distance transit
4 - Ship, boat	1 - Ship | train | coach | airplane	1 – Long distance transit
5 - Van (driver)	2 – Van	2 – Auto (driver)
6 - Van (passenger)	2 - Van	3 – Auto (passenger)
7 - Regional, sub-urban trains	3 – Local public transport	4 – Local public transport
8 - Metro, tram	3 – Local public transport	4 – Local public transport
9 – Bus	3 – Local public transport	4 – Local public transport
10 -Taxi	4 – Taxi	4 – Local public transport
11 - Car (driver)	5 – Car (driver)	2 – Auto (driver)
12 - Car (passenger)	6 – Car (passenger)	3 – Auto (passenger)
13 - Motorbike (driver)	7 – Motorbike	2 – Auto (driver)
14 - Motorbike (passenger)	8 – Motorbike	3 – Auto (passenger)
15 - Moped, scooter	9 – Moped | scooter	2 – Auto (driver)
16 – Bicycle	10 – Bicycle	5 – Bicycle
17 – Walk	11 – Walk	6 – Walk
18 - Other mode	12 – Other mode	4 – Local public transport

The six main modes derived by MiD 2008 are reduced to four for the purpose of the model by:

- omitting trips using long-distance transit modes as daily travel to work does not generally involve long distance travel; and
- combining the categories of auto drivers and passengers into one as this detail is considered irrelevant to the model's purpose of testing the impact of bicycle highways.

Thus, the final choice set for the model includes the following four modes:

1. Walk;
2. Bike;
3. Auto; and
4. Local public transport, referred to hereinafter as 'transit'.

Discarding the trips using long distance transit modes leaves 8,145 trips in the model dataset. Out of these trips, 71% trips are auto trips, 14% are by transit, 10% by bicycle and the remaining 5% by foot. The distances travelled in these trips range from a few meters to approximately 600 km. To better understand the distribution of trips with respect to the distances travelled, the lengths of trips by each of the four modal alternatives were studied. For this, trips with missing trip length information were removed, leaving 8,101 trips in the dataset. The trip length distribution of trips made using each of the four modes is presented in Fig. 5.2.

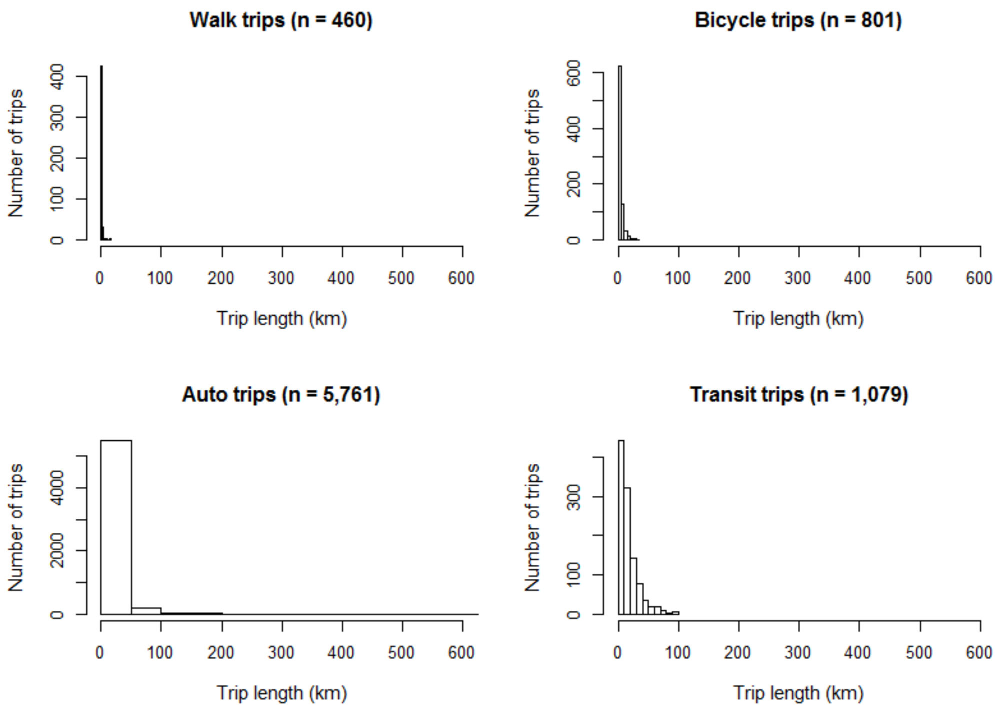

Fig. 5.2: Frequency of trips by trip length for all 4 modes (n = 8,101)

With the exception of auto trips, trips reported by all other modes are under 100 km, and the auto trips beyond this range appear to be outliers. To avoid the risk of affecting the model's accuracy caused by the presence of such outliers, the dataset is limited to trips less than or equal to 100 km. This leaves 8,038 trips in the model dataset. The distribution of mode shares among these trips across travel distances grouped into 5 km bins is shown in Fig. 5.3. Additionally, a frequency distribution of the trips is superimposed for a stronger perspective.

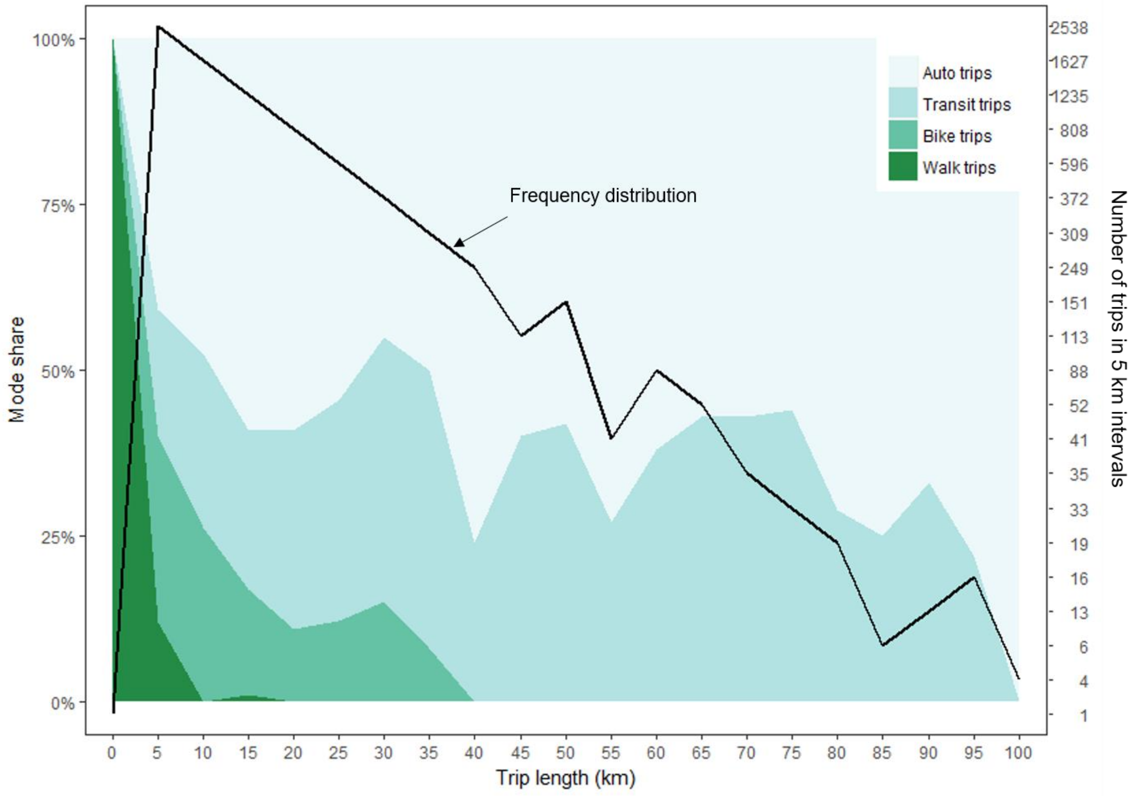

Fig. 5.3: Distribution of mode shares and total trips by trip length[2]

As one would expect, walking and bicycling are dominant for short commutes, and longer commutes are predominantly made by auto and transit. Interestingly, a significant share of bicycle commuting already seems to prevail for trips up to 20 km. The peak beyond 20 km could be a data-inflicted defect as a similar trend can be observed in walk and transit trips as well.

The present aim is to model the individual choice preferences leading to the above mode share patterns. The explanatory variables used to achieve this are discussed in the next section.

5.2.2 Explanatory Variables

As described in section 4.1, choice preferences are measured in terms of utilities of the alternatives, and these utilities are in turn governed by various explanatory variables. The variables used to explain the choice of transport modes are generally classified into the following categories (Ortúzar & Willumsen, 2011, p. 206):

[2] Trips were suitably corrected using weights advised in MiD 2008 data to ensure reliable inferences.

1. Characteristics of the individual, e.g. socio-economic factors, etc.;
2. Characteristics of the journey, e.g. trip purpose, time of the day, number of travel companions etc.; and
3. Characteristics of the transport facility, e.g. travel time, cost, comfort, safety, etc.

While this is a general classification, special classifications emphasizing the importance of variables related to the built-environment (land-use, density, etc.), natural environment (weather, hilliness, etc.) and psychological factors (perceptions, attitudes, etc.) have been proposed for bicycle choice models (Heinen, van Wee, & Maat, 2010; Muñoz, Monzon, & Daziano, 2016). Although such variables better capture individuals' choice to bicycle or not, while modeling choice among multiple alternatives, it may be difficult to gather suitable information through RP surveys. This being the case for the current model and the data available, a general set of variables according to the classification designated by Ortuzar & Willumsen is adopted here.

To select the explanatory variables for the model, those used in similar previous studies reviewed in section 4.2 were consulted. A summary of the variables used in these studies is presented in Appendix A. Based on the review of previous research and the availability of information in MiD 2008 data, the following variables expected to influence commuter mode choice have been selected.

Table 5.2: Explanatory variables considered for the model

Category	**Explanatory variables**
Characteristics of the individual	Age
	Gender
	Income
	Education
	Household size
	Number of children in household
	Auto availability
	Bicycle availability
	Transit subscription
	Distance to nearest transit stop
Characteristics of the journey	Time of day (peak / off-peak)
Characteristics of the transport facility	Travel time

The variables under each category listed above are discussed in detail below.

Characteristics of the individual

The explanatory variables included in this category include socio-economic variables that are most likely to influence the choice of transport mode for trips to work. They include both personal attributes – age, gender, income, education, car availability, bicycle availability, and possession of transit subscription – and household characteristics – household size, number of children in the household, and distance from home to nearest transit stop.

Age and gender

Age and gender are the two most common individual-specific variables considered in mode choice models (See Appendix A).

In the dataset, the age of 9 individuals and the gender of 2 was not reported. Discarding these records with missing age and gender values resulted in a dataset of 8,027 individuals. The resulting dataset comprises of 56% males and 44% females, with ages ranging between 14 and 78. The frequency distribution of male and female commutes in the dataset by their ages is depicted in Fig. 5.4.

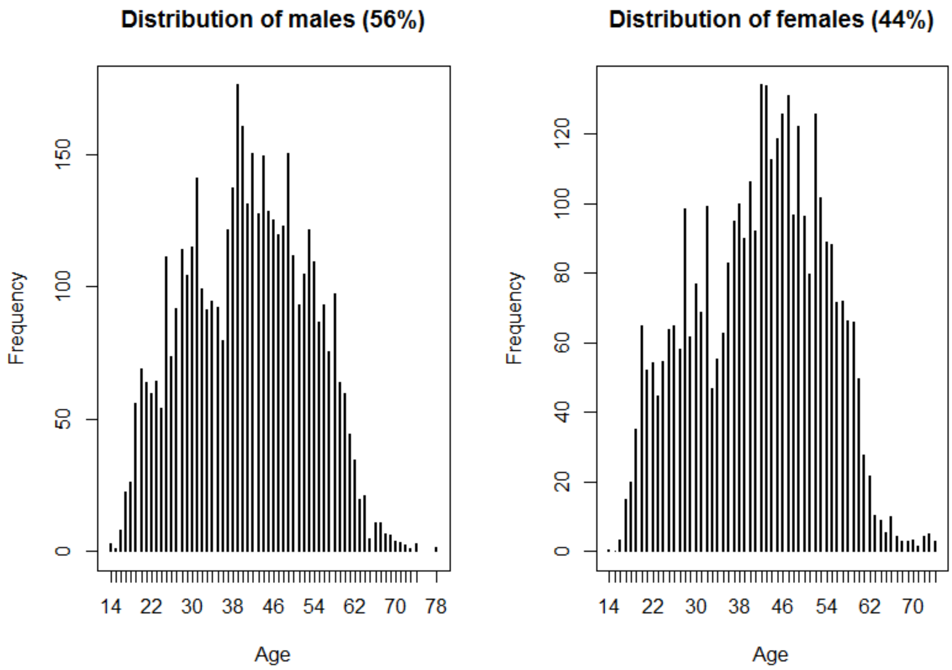

Fig. 5.4: Distribution of male and female commutes by age[3]

[3] Trips were suitably corrected using weights advised in MiD 2008 data to ensure reliable inferences.

As one would expect, a high share of the commuting individuals belongs to the age group of 35 to 60, and the majority of commutes are made by males.

For the model, age was proposed as an integer and gender as a binary variable – 1 for males and 0 for females. These binary variables are called dummy variables, and are useful to see how the behavior of one group within the dataset differs from another (Ben-Akiva & Lerman, 1985, p. 77).

Income

Individuals' choice preferences are greatly influenced by their income budgets, but, 14% individuals in the dataset (1,144) did not disclose income information. Nevertheless, to ensure consideration of income as an explanatory variable in the model, records with missing income information were deleted from the dataset, resulting in 6,883 trips for further analysis.

MiD 2008 records income in terms of net monthly household income grouped as follows:

1 - €500 or less
2 - €500 to < €900
3 - €900 to < €1,500
4 - €1,500 to < €2,000
5 - €2,000 to < €2,600
6 - €2,600 to < €3,000
7 - €3,000 to < €3,600
8 - €3,600 to < €4,000
9 - €4,000 to < €4,600
10 - €4,600 to < €5,000
11 - €5,000 to < €5,600
12 - €5,600 to < €6,000
13 - €6,000 to < €6,600
14 - €6,600 to < €7,000
15 - Greater than €7,000

As mode choice behavior among individuals differs across income groups, mode shares of individuals in the dataset were analyzed. The distribution of mode shares across income groups is shown in Fig. 5.5. The frequency of trips across the different income groups is also included for a better perspective.

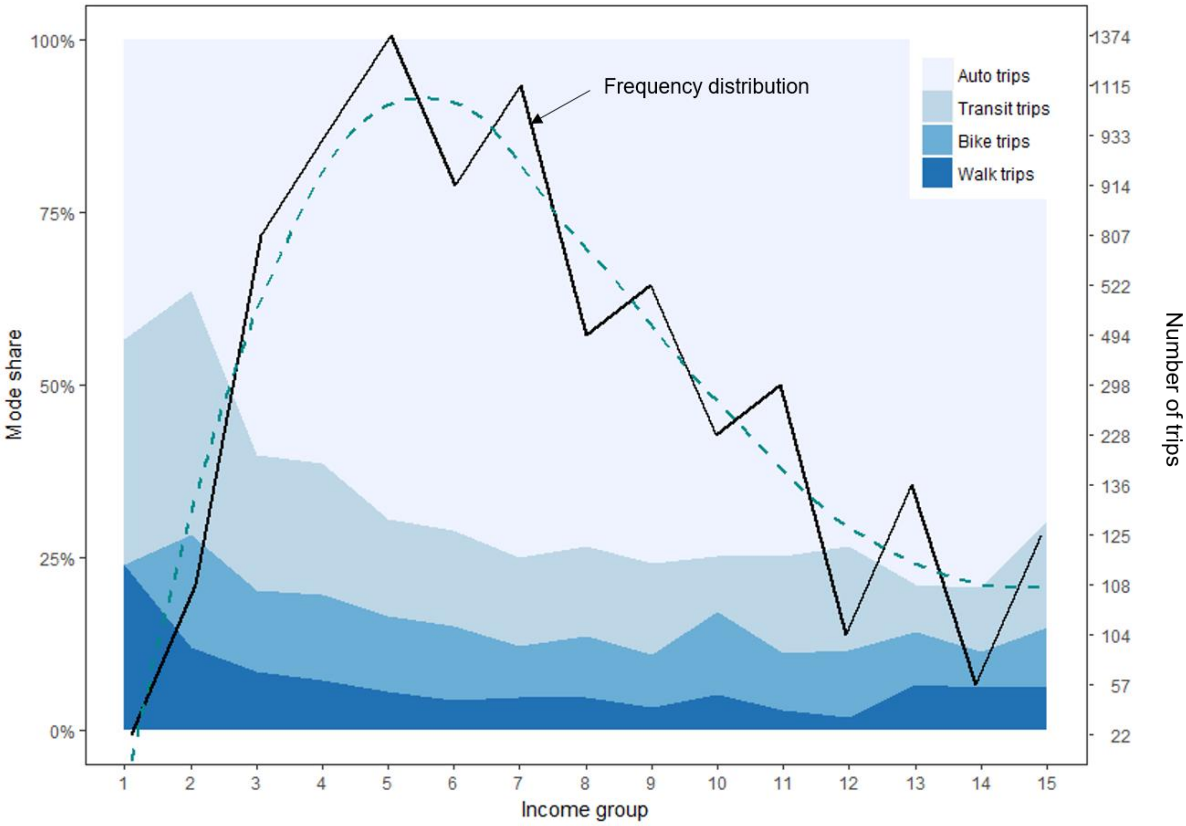

Fig. 5.5: Distribution of mode shares and total trips by income group[4]

Evidently, transit, walk and bike shares are dominant for lower income groups, and they gradually decline as income increases. However, an increase in the shares of these modes can be seen between income groups 14 and 15, which is perhaps due to inherent data deficits. Further, very few commute trips are reported for low and high income groups. Consequently, no reliable variation in mode choice trends can be observed across the different income groups, and hence the dataset could not be split based on income groups to run different estimates for the individuals from different income classes. Instead, income was proposed to be considered as a continuous variable. To do this, continuous income values were engineered based on the frequency distribution of trips across the reported income groups.

The frequency distribution of trips across income groups can be seen to follow a gamma distribution (see the smoothened distribution represented by a green dashed line in Fig. 5.5), as is the case in many transport modelling studies (MacKinder, Evans, & May, 1975; Wootton & Pick, 1967). Considering the mid-range values of income groups, a gamma distribution with a shape factor 2.79 and an inverse rate factor of 850.82 was found to fit the dataset most

[4] Trips were suitably corrected using weights advised in MiD 2008 data to ensure reliable inferences.

closely. By fitting this gamma distribution to the distribution of income groups, each record in the dataset was assigned a random income value within the range of her or his income group. As the values so assigned correspond to household incomes, they were divided by the number of working individuals in the household to obtain individual incomes. A limitation of this approximation is that all working individuals from one household are assigned the same income.

Education

MiD 2008 data reports the highest education level obtained by individuals. To analyze differences in choice behavior among individuals with college education and those without, a dummy variable indicating individuals' attainment of a college degree was proposed for consideration in the model (1 for individuals with a college degree and 0 for those without). However, as a high share of individuals in the dataset (43%) did not disclose this information, this variable was dropped.

Household size and number of children

Household attributes like household size, number of children and number of employed people have an influence on individuals' choice of mode. For instance, household members may combine or chain individual trips, making auto travel convenient. To capture this effect, these variables were proposed for consideration in the model.

MiD 2008 data reports the number, age and employment status of individuals in the households. In the dataset, one individual's household attributes were not revealed, and hence that record was dropped. In the remaining 6,882 records, individual household sizes range from 1 to 7, with both number of children and number of employed persons varying between 0 and 6. Here, individuals under 18 years of age were considered as children, considering the minimum age requirement for unrestricted driving in Germany (Bundesministerium für Verkehr und digitale Infrastruktur, 2016) as they may require adults to drive them.

Auto availability

Having access to an auto significantly increases one's propensity to drive to work. This propensity is further increased when an individual has access to more than one vehicle. To capture this, auto availability for individuals' in the dataset was proposed to be evaluated as the number of autos available to the individual calculated as given in Equation 5.1.

Equation 5.1

$$Auto\ availability = \frac{Number\ of\ household\ autos}{Number\ of\ household\ members\ posessing\ driver's\ license}$$

MiD 2008 data records both the number of autos and the number of persons in households. However, as the households of 2 individuals in the dataset did not report the number of autos, these records were deleted. In the resulting dataset of 6,880 records, the number of household autos and driver's licenses range from 0 to 7 and 0 to 6 respectively. The auto availability calculated as given in Equation 5.1 thus varies between 0 and 7. However, the value 7 is an outlier as the majority of individuals in the dataset have up to 2 autos available (see Fig. 5.6).

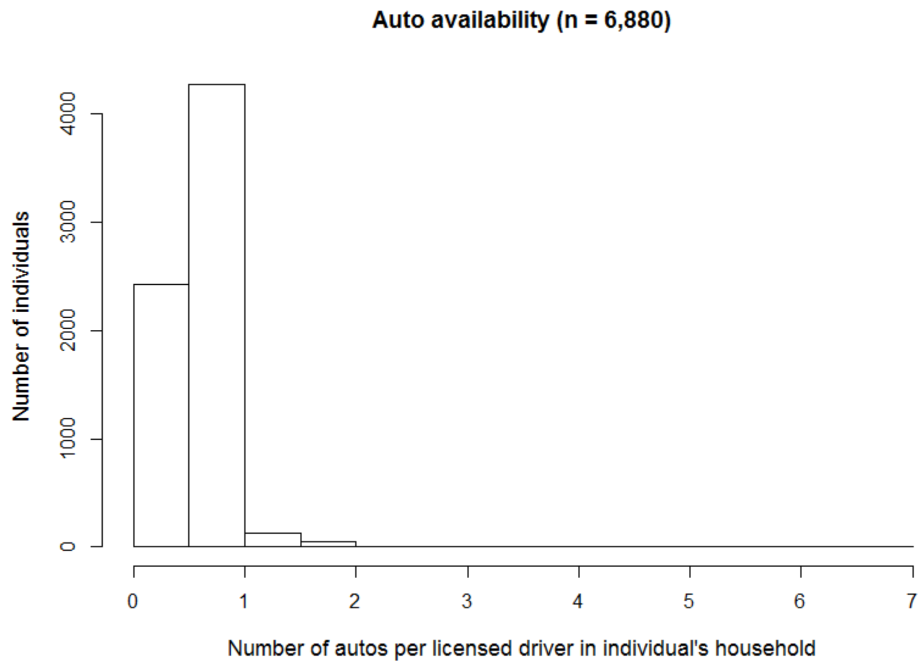

Fig. 5.6: Auto availability of individuals in the dataset

Bicycle availability

Access to a bicycle naturally increases one's chance of choosing to ride. Consequently, a dummy variable indicating an individual's bicycle availability is proposed for the model (1 for individuals with an available bicycle, and 0 for those without).

Individuals' surveyed in MiD 2008 were asked to report whether or not they owned a functional bicycle, indicating their bicycle availability. Out of the records in the dataset, 29% (1,987) of individuals did not reveal this information. However, the survey also records information about the number of bicycles in a household. All but one of the 29% of records report this information. By eliminating this record with no information on both access to bicycle and number of household bicycles, bicycle availability for the 29% was engineered from the number of household bicycles and the number of persons in the household.

In the resulting dataset with 6,879 individual trips, bicycle availability was assigned as the individual's access to bicycle where this information was available, and for the remaining individuals (29%), bicycle availability was allocated according to the rule given in Equation 5.2.

Equation 5.2

$$Bicycle\ availability = 1\ if\ \frac{Number\ of\ household\ bicycles}{Nubmer\ of\ household\ members} \geq 1;\ \ and\ 0\ otherwise$$

Consequently, 79% of individuals in the dataset were assigned to have a bicycle available. The validity of this derivation can be evaluated against the typical bicycle ownership in the country. On an average, 68.5% Germans own at least one bicycle (pressedienst-fahrrad GmbH, 2014). Although 79% is on the higher side, it is to be noted that the national average includes individuals from both urban and rural areas while the model dataset includes only employed individuals in urban areas, and rural areas generally have a lower bicycle ridership compared to urban areas.

Transit subscription

Regular transit users often subscribe to a transit pass. The possession of a transit subscription therefore indicates a preference for transit use. To capture this, considering a dummy variable indicating whether an individual is subscribed to a transit pass or not (1 – Yes, 0 – No) was proposed. MiD 2008 data reports the type of transit ticket predominantly used by individuals. However, as 40% (2,735) of individuals in the dataset did not reveal this information, the variable had to be dropped.

Distance to nearest transit stop

Transit proximity is a major factor that governs the choice of transit for a trip. When no transit stop is available within a convenient distance from an individual's house, the chance of her or him traveling by transit naturally reduces. Although this information was reported in the survey data, it was unavailable for many individual households in the dataset (30% = 2,006 individuals). Without information on the location of the households, it was not possible to approximate the values, and hence the variable was discarded.

Characteristics of the journey

Variables that fall under this category include the characteristics of the trip itself that answer questions like when, why, with whom etc., the trip is made. Trip purpose, time of day when the trip is made, how many companions was the trips made with, etc. fall under this category. Among these, the time of day is considered here as trip purpose has been accounted for by considering only commutes for the study, and travel companion attributes are not deemed relevant for the model's purpose.

Time of day (peak / off-peak)

The choice of mode is expected to be affected by the time of day, especially by traffic peak hours. Given an opportunity, travelers would try to avoid trips during peak hours to escape congestion on roads and crowds in transit. Further, bicycle highways are being promoted as an attractive option for commuters mainly to avoid peak hour travel issues. Their effect could be observed in the proposed model application with the help of a dummy variable differentiating peak-hour trips from off-peak hour trips.

For this, trip start times given in the survey data grouped as follows were used -

1 – before 05:00
2 – 05:00 to 07:00
3 – 07:00 to 09:00
4 – 09:00 to 11:00
5 – 11:00 to 14:00
6 – 14:00 to 17:00
7 – 17:00 to 20:00
8 – 20:00 to 22:00
9 – after 22:00

Trips made in time groups 3 (07:00 to 09:00) and 7 (17:00 to 20:00) were considered as peak-hour trips, represented by 1, and trips starting in all other time groups were marked 0.

Characteristics of the transport facility

This category comprises of variables related to the transport modes and facilities. Time and cost attributes are the most commonly modeled variables under this category. Another variable often considered under this category is the number of transfers necessary for travel by transit. However, travel costs were not recorded in MiD 2008 data. Had there been information on home and work locations of individuals in the data, costs could have been engineered from other sources. As this was not possible, only travel time was considered for the model.

Travel time

Travel time is generally measured in terms of its components – access time, in-vehicle time, and egress time. Additionally, waiting time and number of transfers are included for transit trips. However, this level of detail is unavailable in MiD 2008 data. The survey only records start and end times of individuals' trips, indicating the total trip duration. This duration is considered as the travel time for the model.

Furthermore, the survey data only reports the travel times corresponding to the modes used by the individuals. Travel times for modes that were not chosen, therefore, had to be engineered. Ideally, these times should be estimated through a route choice assignment exercise, as done by Halldórsdóttir et.al. (2011). But, due to the lack of locational information, the times were engineered from travel distances and average speeds instead. In the MiD 2008 data, trip distances were reported by individuals, and the speeds were calculated from the reported travel times and distances and corrected for implausible entries. Although the distances correspond to the mode chosen, they were assumed to be the same across modes due to a lack of information on the origin and destination locations. To estimate the travel times, first average speed of the reported trips by different modes was calculated incorporating sample weights. Before doing this, trips in the dataset with blanks for travel times (2) and speeds (15) were removed, resulting in a dataset of 6,862 trip records. The calculated average speeds of trips by the different modes are indicated in Table 5.3.

Table 5.3: Average speeds of trips by travel mode

Mode	**N***	**Average speed (km/h)***
Walk	398	4.8
Bicycle	730	13.7
Transit	1,079	20.9
Auto	5,013	38.6

* - weighted values

The above speeds were used to determine travel times for modes not chosen from the travel distances reported.

Before finalizing the set of explanatory variables to be applied for the estimation of the model, variable correlations were verified to ensure that no two variables are correlated, as discrete choice theory requires the variables to be mutually independent. The correlations among all the above variables are shown in Fig. 5.7.

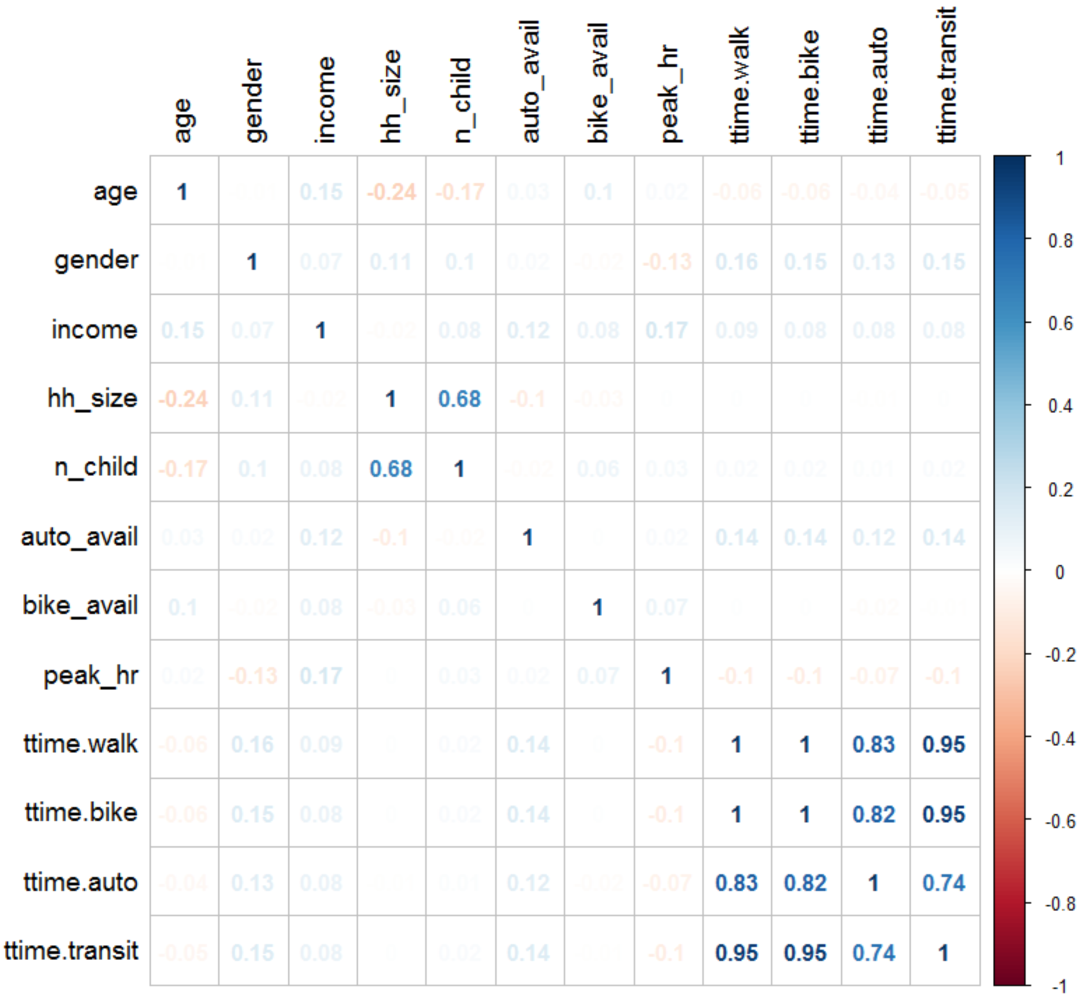

Fig. 5.7: Variable correlations

Fig. 5.7 shows a significant correlation between household size and the number of children. To avoid discrepancies caused by such a correlation, estimations involving these variables were carefully examined. No other variables seem to exhibit major correlations. Travel time is one explanatory variable, hence the correlations among travel times of the different alternatives do not imply any harm to the model's framework. After all, they are of the same metric and those not reported were estimated based on average speeds.

An overview of the list of explanatory variables adopted for model estimation based on the final dataset of 6,862 trips is given in Table 5.4 below –

Table 5.4: Overview of explanatory variables in the model dataset (n = 6,862)

Variable	Range	Mean
Age	14 – 78	44
Gender (1 – Male, 0 – Female)	-	-
Income	€118 - €9,204	€1,847
Household size	1 – 8	3
Number of children	0 – 6	0.6
Auto availability (No. autos per driver's licenses)	0 – 7	0.8
Bicycle availability (1 – Yes, 0 – No)	-	-
Time of day (1 - Peak / 0 - Off-peak)	-	-
Travel time by walk (min)	1 – 75	12
Travel time by bicycle (min)	2 – 90	17
Travel time by transit (min)	1 – 475	46
Travel time by auto (min)	1 – 480	25

Due to data limitations, level of service attributes beyond those in the above table could not be considered for estimation. Although this would decrease the efficiency of the statistical estimates, the information that could not be included like travel cost and individual time components is not expected to impede the model's purpose of determining the impacts of bicycle highways. However, availability of information on origin (home) and destination (work) locations of the commutes would have significantly enhanced the range and quality of the model attributes.

5.2.3 Model Structure

With the availability of the dataset comprising information on individual commute trips, a disaggregate discrete choice model was adopted. Under this framework, choice probabilities of alternatives are established by maximizing utilities of each alternative for the target group of individuals. Subsequent to the last two sections where the choice set and the explanatory variables were determined, utilities of the alternatives involved and the modelling structure to be adopted are established in this section.

According to the choice modeling theory described in section 4.1, utilities are comprised of a measureable systematic component and a random error component. Systematic components are expressed as additive linear-in-parameter functions of known explanatory variables and unknown parameters, and random components are assumed to follow a certain probability distribution. Under this framework, the utilities of the four alternatives of the model's choice set are written as follows –

Equation 5.3

$$U_{walk} = V_{walk} + \varepsilon_{walk}$$

$$U_{bike} = V_{bike} + \varepsilon_{bike}$$

$$U_{transit} = V_{transit} + \varepsilon_{transit}$$

$$U_{auto} = V_{auto} + \varepsilon_{auto}$$

The corresponding systematic utilities of the alternatives expressed as the sum product of the explanatory variables and unknown β parameters are formulated as follows –

Equation 5.4

$$\begin{aligned} U_{walk} = {} & \beta_{0w} + \beta_{1w} * age + \beta_{2w} * gender + \beta_{3w} * income + \beta_{4w} * household\ size + \\ & \beta_{5w} * number\ of\ children + \beta_{6w} * number\ of\ workers + \beta_{7w} * auto\ availability + \\ & \beta_{8w} * bicycle\ availability + \beta_{9w} * time\ of\ day + \beta_{10w} * travel\ time\ by\ walk \end{aligned}$$

$$\begin{aligned} U_{bike} = {} & \beta_{0b} + \beta_{1b} * age + \beta_{2b} * gender + \beta_{3b} * income + \beta_{4b} * household\ size + \\ & \beta_{5b} * number\ of\ children + \beta_{6b} * number\ of\ workers + \beta_{7b} * auto\ availability + \\ & \beta_{8b} * bicycle\ availability + \beta_{9b} * time\ of\ day + \beta_{10b} * travel\ time\ by\ bicycle \end{aligned}$$

$$\begin{aligned} U_{transit} = {} & \beta_{0t} + \beta_{1t} * age + \beta_{2t} * gender + \beta_{3t} * income + \beta_{4t} * household\ size + \\ & \beta_{5t} * number\ of\ children + \beta_{6t} * number\ of\ workers + \beta_{7t} * auto\ availability + \\ & \beta_{8t} * bicycle\ availability + \beta_{9t} * time\ of\ day + \beta_{10t} * travel\ time\ by\ transit \end{aligned}$$

$$\begin{aligned} U_{auto} = {} & 0 + \beta_{1a} * age + \beta_{2a} * gender + \beta_{3a} * income + \beta_{4a} * household\ size + \\ & \beta_{5a} * number\ of\ children + \beta_{6a} * number\ of\ workers + \beta_{7a} * auto\ availability + \\ & \beta_{8a} * bicycle\ availability + \beta_{9a} * time\ of\ day + \beta_{10a} * travel\ time\ by\ auto \end{aligned}$$

Where, $\beta_{0w}, \beta_{0b}, \beta_{0t}$ are the alternative specific constants,
$\beta_{1w}, \beta_{2w}, \ldots, \beta_{8w}$ are the unknown β parameters corresponding to walking utility,
$\beta_{1b}, \beta_{2b}, \ldots, \beta_{8b}$ are the unknown β parameters corresponding to bicycling utility,
$\beta_{1t}, \beta_{2t}, \ldots, \beta_{8t}$ are the unknown β parameters corresponding to transit utility, and
$\beta_{1a}, \beta_{2a}, \ldots, \beta_{8a}$ are the unknown β parameters corresponding to auto utility.

In the above utility equations, the alternative auto was considered as the base. As utility is a relative measure, the parameters are estimated relative to one of the alternatives known as the base. Further, as constants do not vary across outcomes, the constant of the base alternative is set to 0 (Washington, Karlaftis, & Mannering, 2011).

With the above utilities, the current model assumes a logit modelling framework where the random component is assumed to follow a Gumbel distribution as described in section 4.1. Most similar studies predominantly use either MNL or NL frameworks (refer Appendix A). While

MNL is the most popular structure due to its many advantages including straightforward formulation, NL is used when correlations between alternatives are expected as the IIA property of MNL does not allow for this. In this thesis, both MNL and NL model estimations were attempted as no particular evidence of the suitability of either structure was detected. The estimations and results of the resulting model are discussed in the next section.

5.3 Model Estimation and Results[5]

This section discusses the estimation of the unknown β parameters included in the model's utility functions as illustrated in section 5.2. As described in section 4.1, such discrete choice models are estimated through maximum log likelihood estimations. In this thesis, logit class estimations – both MNL and NL – were estimated using the statistical computing software R as such models with multiple alternatives and explanatory variables involve numerous iterations, making manual estimations almost impossible. R was used in this thesis because it has a very comprehensive framework, can handle complex statistical procedures through its numerous packages, and is freely accessible (Viton, 2015). This thesis used R's 'mlogit' package which contains functions enabling logit estimations. The estimation procedure and results are described in the following subsections.

5.3.1 Estimation Using *mlogit* R Package

The mlogit R package was developed by Yves Croissant for the estimation of MNL models and can be extended to NL among others (Croissant, n.d.[2011]). A detailed description of the application of the package can be found in Croissant, n.d.[2011] and Viton, 2015. Only the aspects of the package used in this thesis are discussed here.

The procedure involves two main steps –
1. specifying the dataset; and
2. calling a function that performs the estimation.

Dataset specification

The mlogit package accepts data in two formats – wide shape and long shape. A wide dataset has one row for each observed choice; whereas in the long shape arrangement, there is one row for each alterative of every choice situation. Wide shape is preferred when the model involves many individual specific explanatory variables (characteristics of trip makers) as these

[5] The estimations and analysis discussed in this section were performed based on the records of the dataset duly corrected using the sample weights indicated in the survey data.

do not vary across alternatives, and in cases where most of the variables are alternative specific (characteristics of transport facilities), long datasets are preferred.

In the current model, as most of the explanatory variables are individual specific (see Table 5.4), a wide format was adopted. A sample of the dataset arranged in the wide format is shown in Appendix B.

The so arranged dataset was supplied to R using the function *mlogit.data()* where the shape (wide / long), the variable indicating the chosen alternative, and the indices of the alternative-specific variables (e.g. travel time) were explicitly indicated. The specified data was then used to call the function for estimation.

Function call for estimation

The mlogit package has a built-in function called *mlogit()* for performing logit estimations. The structure of the model to be estimated is specified to the function by indicating how explanatory variables are to be considered. Three types of explanatory variable specifications can be entered into *mlogit()* –

1. Alternative specific variables;
2. Individual specific variables; and
3. Alternative specific variables with coefficients (β parameters) differing across alternatives.

An MNL estimation using *mlogit()* with variable types described above is called as follows –

Equation 5.5

$$mlogit(\,choice \sim \,Type\ 1\ variables\ \mid Type\ 2\ variables\ \mid Type\ 3\ variables, data = \ dataset)$$

In the current model, all explanatory variables except travel times were considered as individual specific (type 2), and travel times were considered under type 3 assuming that changes in travel time affect different modes differently. For instance, travelers may be more accepting of increased travel times for driving to work than they would be for walking.

The formulation of NL estimation is very similar to that of MNL. Alternatives and datasets are specified just as in Equation 5.5 with only an additional specification of the nesting structure. This thesis attempted a nesting structure containing two nests – 'non-motorized' with walk and bicycle alternatives, and 'motorized' with auto and transit alternatives. The function call for this structure in comparison to the MNL one in Equation 5.5 can be written as –

Equation 5.6

$$mlogit(\ choice \sim\ Type\ 1\ variables\ \mid Type\ 2\ variables \mid Type\ 3\ variables, data = dataset,$$
$$nests = list(non\_motorized\ =\ c("walk", "bike"), motorized\ = c("auto", "transit"))$$

With the help of the functions described above, MNL and NL estimations were performed to identify the set of explanatory variables that best fit the choice behavior of the individuals in the dataset. A discussion of the results of the estimations is presented in the following section.

5.3.2 Results and Discussion

The model was first estimated for a MNL framework. A series of estimations were performed for different combinations and forms of the set of explanatory variables. This was done by first considering all variables in the model specification as per Equation 5.5. Variables which were found to be statistically insignificant or theoretically inconsistent were then removed one after another to obtain a model fit with at least 95% significance. The main findings of this exercise are listed below:

- The dummy variable ‘time of day’ indicating peak-hour trips was statistically insignificant for all four modes. This may be explained because most commutes are generally made during certain hours regardless of the traffic situation owing to fixed working times.
- Bike availability was found to be irrelevant in explaining the choice of transit for work trips, and hence was removed.
- Males were predicted to walk to work more than females, but the results were not statistically significant.
- Most likely due to the correlation between household size and number of children, the presence of one made the other statistically insignificant. As household size encompasses the number of children, the variable number of children was deleted from the estimations.
- Surprisingly, the variable for personal income was not statistically significant for any of the modes, and was removed. Different functional forms – logarithmic, quadratic and exponential – forms of the variable were attempted, but none improved its significance. A similar result was encountered by Halldórsdóttir, Christensen, Jensen, and Prato (2011) while estimating their model of mode choice for trips less than 22 km in Denmark. They also highlight the contrasting results found in different studies for the impact of income on bicycle shares.
- Travel time was insignificant for transit commutes in its simple form. It is to be noted that due to data limitations, the travel times considered in the model were total trip durations,

while most mode choice models consider in-vehicle, access and egress times as separate variables, especially for transit (Halldórsdóttir, Christensen, Jensen, & Prato, 2011; Ortúzar, Iacobelli, & Valeze, 2000). In this regard, it is difficult to comment on this result. However, a quadratic form resulted in a high significance level, perhaps because it mitigates the effect of time rises.

Subsequent to the above exercise, an estimation where all the variables explain the model at a minimum significance level of 95% was obtained. The significance level was measured by a t-test where a null hypothesis that the corresponding parameter is zero is tested (Ben-Akiva & Lerman, 1985, p. 93). A summary of the results of the estimation are shown in Table 5.5.

Table 5.5: MNL estimation results

Variable	Estimate	t-value	Significance
Constant - Auto	0.0000		
Constant - Walk	5.6753	12.101	***
Constant - Bicycle	1.2523	3.806	***
Constant - Transit	2.2955	10.023	***
Age - Walk	-0.0134	-2.125	*
Age - Bicycle	-0.0132	-3.164	**
Age - Transit	-0.0208	-6.307	***
Gender - Bicycle	0.2438	2.639	**
Gender - Transit	-0.3274	-4.099	***
Household size - Walk	-0.3483	-5.467	***
Household size - Bicycle	-0.0859	-2.090	*
Household size - Transit	-0.2415	-6.762	***
Auto availability - Walk	-2.2565	-9.854	***
Auto availability - Bicycle	-3.1013	-19.358	***
Auto availability - Transit	-3.4413	-25.589	***
Bicycle availability - Walk	-0.4251	-2.429	*
Bicycle availability - Bicycle	1.9764	10.339	***
Travel time - Walk	-0.1523	-19.136	***
Travel time - Bicycle	-0.0749	-21.423	***
Travel time - Auto	-0.0186	-7.124	***
(Travel time) ^{2} - Transit	-0.0001	-5.670	***

*** 99.9% significance level, ** 99% significance level, * 95% significance level

Table 5.5 shows the estimates of the parameters for the variables included in the final model estimation. It also included the results of the t-tests – t-values and levels of significance. A statistical summary of the estimation is given in Table 5.6.

Table 5.6: Statistical summary of the MNL estimation

Variable	Estimate
Number of observations	6,862
Number of iterations	208
Log-likelihood	-4113.2
McFadden R^2	0.34603
Likelihood ration test: chisq	4352.7 ($p < 2.22e-16$)

The complete result output of the estimation from R is shown in Appendix C for further reference. The estimation was performed with the mode auto as the reference as it is the most represented. The estimates are hence to be understood with respect to auto, the base mode. A discussion of the results is as follows:

Travel time

The travel time parameter estimates of all four modes are consistently negative. This is expected as travel times increase, utility decreases. Further, as one would imagine, the estimate for travel time by walk is the most negative, followed by bicycle and auto; implying that for an increase in travel time to work, walking becomes the least attractive mode, and bicycling the next least attractive and auto the next. Transit travel time is estimated to have the least negative parameter indicating that transit becomes the least unattractive of all modes for an increase in travel time to work. This could imply that the individuals who reported transit commutes in the dataset are predominantly captive riders. Another explanation could be that the majority of employment opportunities are often found in urban centers and due to reasons like congestion and parking restrictions, commuters travelling to city centers prefer transit. The quadratic function, however, perhaps also contributed to this.

Bicycle availability

As one would expect, bicycle availability has a positive effect on the propensity of bicycling to work. The tendency to walk to work is further reasonably seen to decrease when there is a bicycle available.

Auto availability

Logically, the availability of an auto reduces one's tendency to use the other modes. The model emulates this with the negative parameter estimates for walk, bicycle and transit. Further, auto availability seems to affect transit ridership most strongly, followed by bicycle and walk in that order.

Household size

The results indicate that commuters from larger households are less likely to walk, bicycle or take transit to work than they are to drive an auto. The effect appears to be the strongest on walking and transit. This is perhaps because individuals from larger households often tend to combine trips to accompany other household members, or chain other trips like shopping or dropping and picking up children along with their commutes, which makes auto travel more convenient.

Gender

As the value 1 of the dummy variable gender was assigned to males, the results directly relate to the group of male commuters. The positive parameter estimate for bicycling indicates that males are more like to bike to work than females, a result that has been observed in many studies (Halldórsdóttir, Christensen, Jensen, & Prato, 2011; Rodríguez & Joo, 2004). The estimation further indicates that males are less likely to take transit to work than females, perhaps because they are usually found to bicycle and drive more.

Age

The estimation logically shows that as individuals get older, they are less likely to walk or bike or take transit to work than they are to drive. The impact, however, seems to be the strongest on transit. The impact on walk and bike appears similar.

Mode-specific constants

The mode-specific constants shown in the results capture the effects of the unobserved variables and measurement errors. Auto has a constant value of 0 because it is the base mode. The other three modes have positive constants relative to auto, indicating that all else being equal, auto is the least likely to be chosen, followed by bicycle and transit. As expected, walking takes the highest constant as it is the least represented in the dataset, and its attributes were perhaps not well captured by the model. The absence of transit attributes also justifies the magnitude of its constant.

To estimate the accuracy of the model fit, choice probabilities were predicted for the same model dataset using the estimated parameters. The difference in the mean share of the four modes obtained with respect to the shares in the dataset is indicated in Table 5.7.

Table 5.7: Result of a test for accuracy of model fit

Mode	Actual share	Predicted share	Deviation
Walk	5.5%	5.6%	0.1%
Bicycle	10.1%	10.4%	0.3%
Transit	14.9%	13.3%	-1.6%
Auto	69.4%	70.7%	1.3%

The deviation of the mean mode shares predicted by the model compared to the actual shares of the dataset were found to be within a range of $\pm$2%.

NL estimations for the nesting structure described in section 5.3.1 were also attempted, but a reasonable nesting coefficient could not be estimated using the mlogit R package. The obtained coefficients were greater than 1, and hence the model had to be rejected (Koppelman and Bhat 2006, p.163). This can be explained as a data limitation as the available information was too coarse to estimate correlations between alternatives. With a finer data resolution, for instance with information on home and work locations, a broader range of variables could have been considered, allowing a better evaluation of interactions between alternatives. As NL models could not be estimated reasonably, the MNL estimation described above was applied in the scenario analysis described in the next chapter.

6 Scenario Analysis

The objective of this thesis is to assess the impact of the bicycle highways proposed in the Munich region on commuter mode share. For this assessment, a commuter mode choice model was built based on urban commute data from the MiD 2008 survey as described in chapter 5. This model was then applied to the scenario of the pilot bicycle highway recommended in Munich. The results of this application of the model to the scenario of a bicycle highway are described in this chapter.

Chapter 2 gave a brief account of the bicycle highways planned in Munich. As stated, the city's pilot bicycle highway is proposed to run between the city center and a northern suburb, Garching (Planungsverband Äußerer Wirtschaftsraum München, 2015b). Fig. 6.1 marks the highway chosen as the pilot by experts from the six shortlisted routes.

Figure removed due to possible copyright infringements

Fig. 6.1: Munich's pilot bicycle highway (Adapted from (Planungsverband Äußerer Wirtschaftsraum München, 2015a, p. 51))

This pilot route was selected for a scenario analysis in this thesis. By applying the model parameters estimated in section 5.3 to a dataset of commuters who would have an option of using this bicycle highway in a hypothetical scenario, their choice probabilities were predicted. The setup of the dataset and the results of the prediction are discussed in the following sections.

6.1 Dataset

For the analysis, a dataset of individuals with home and work locations along the proposed highway corridor was prepared. Data for this purpose was borrowed from a synthetic population generated by the Modeling Spatial Mobility research group for their work on developing an integrated land-use and transportation model for the Munich region. Information on the creation of the synthetic population can be found at Technical University Munich, 2016.

Based on a geographical zoning system developed by the research group, individuals living and working along the corridor were identified by superimposing the corridor on the zones. Information about individuals with home and work locations in the zones located along the corridor, indicated in green in Fig. 6.2, was then extracted from the synthetic population.

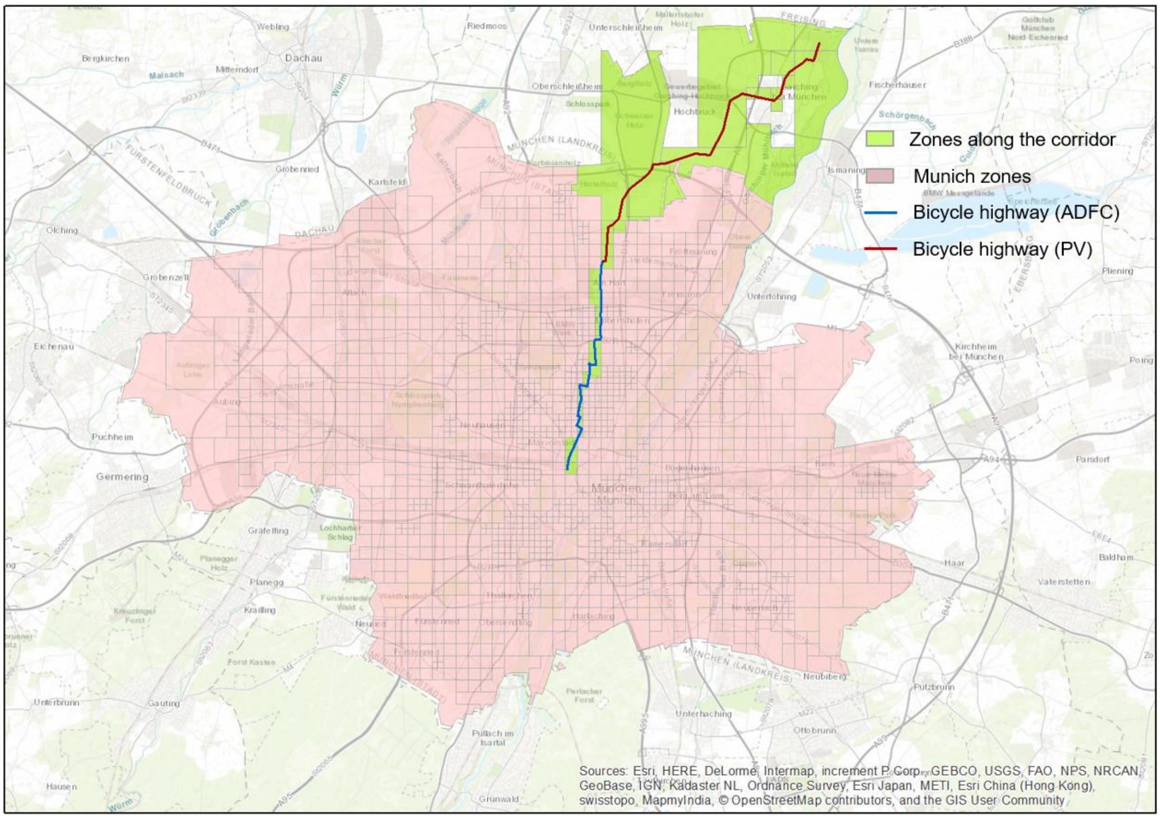

Fig. 6.2: Selection of zones along the bicycle corridor based on a zoning system

The bicycle highway corridor, indicated in blue in Fig. 6.2, was plotted in ArcGIS based on the detailed route plan published by Munich's Bicycling Association (ADFC München e.V., 2014). As this route did not extend until Garching, the route was completed (in red) based on the Planning Association (PV)'s tentative plan (Fig. 6.1). The selected zones yielded a dataset of 450 commuters who could potentially use the corridor. Further commuters could not be considered due to the model's limitations. If it were possible to combine the model with a route choice component, as done by Halldórsdóttir et al. (2011), a detailed analysis of prospective commuting users of the corridor could have been performed instead.

The synthetic population provided age, gender, number of children, household size and income of the individuals in the dataset. As information on bicycle availability and auto availability was unavailable, these variables were estimated through regressions from the available attributes.

Bicycle availability being a categorical variable (see Table 5.4), a logistic regression was performed in R to estimate the same for the individuals in the scenario dataset. A random sample of the model dataset discussed in chapter 5 was used to train and test the regression. The best fit was obtained for the variables age, household size, number of children and a logarithmic transformation of income. The summary of a resulting estimation is shown in Appendix D. When tested on random samples from the model dataset, the regression predicted bicycle availability with an average accuracy of 79.9%. Although the regression was not highly accurate, due to lack of a better source, it was applied to estimate the bike availability of the individuals in the scenario dataset. Subsequently, 98% of the 450 individuals in the scenario dataset were assigned to have a bike available.

In the model, auto availability was measured as the number of autos per licensed driver in the household. As this is a continuous variable, a multiple regression was performed to estimate values based on the available attributes of age, gender, household size, number of children and income. The procedure was similar to that adopted for the estimation of bicycle availability. Regressions were performed on the model dataset, and the best fit, which was observed with gender, household size, number of children and a logarithmic transformation of income, was in turn tested on the model dataset. The summary statistics of the actual and predicted auto availability values obtained from the regression are shown in Table 6.1.

Table 6.1: Comparison of actual and predicted auto availability of the model dataset

Auto availability	**Minimum**	**Mean**	**Maximum**
Actual	0	0.7654	7
Predicted	0.4843	0.7686	1.003

Although the ranges differ, the predicted mean values are sufficiently close. However, as described through Fig. 5.6 in section 5.2.2, the actual auto availability values contain outliers. This can also be seen in the regression fit depicted in Fig. 6.3 below.

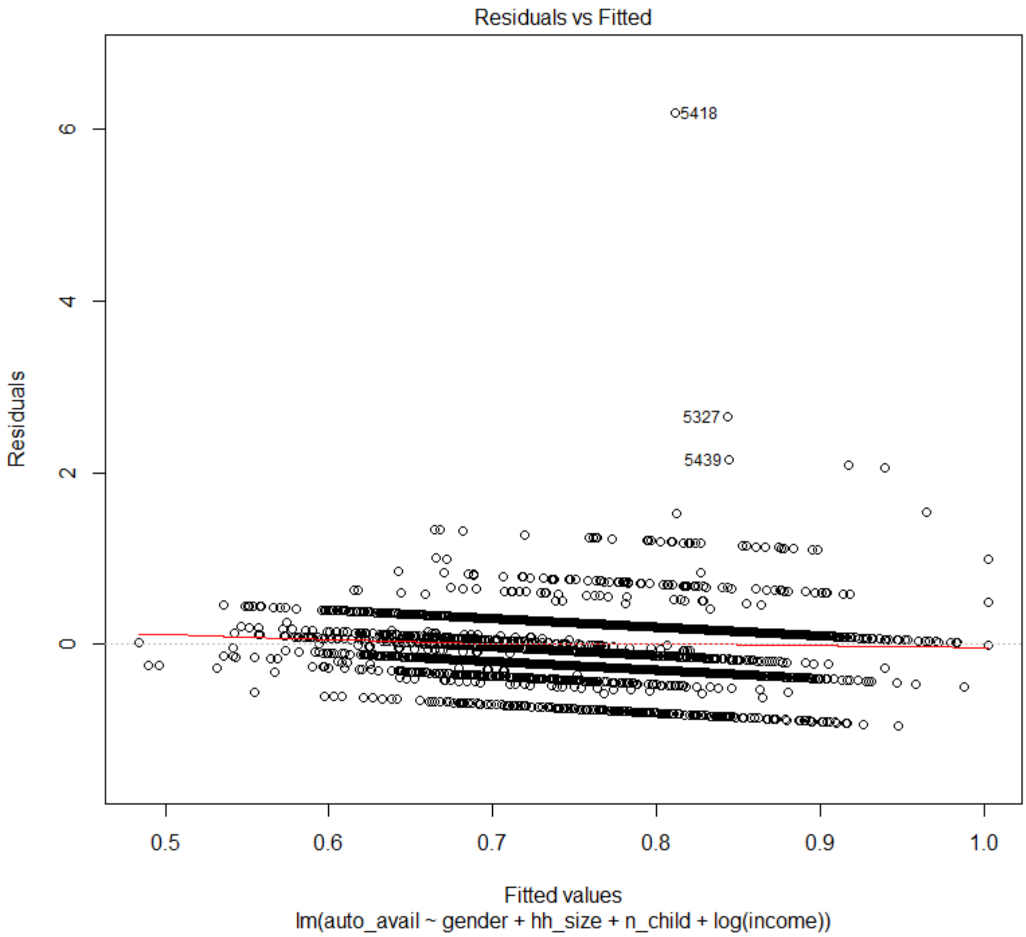

Fig. 6.3: Regression fit for auto availability

The fit resulted in a root mean squared error of 0.35 and a mean absolute error of 0.26. The summary of the regression output can be found in Appendix E.

In addition to the above socio-economic variables, the scenario dataset required the individual's commute travel times to enable predictions using the model. To obtain travel times, first, centroid coordinates of all the chosen zones along the corridor were established using ArcGIS. Then, a Java script for extracting travel times between point coordinates from Google Maps, which was developed by the Modeling Spatial Mobility research group, was used to extract travel times and distances by all four modes. The extraction was done for commutes on 07.02.2017, a random Tuesday, at 8:00 am, to ensure a date not too far ahead in the future. Finally, based on the home and work zone information of the individuals available in the synthetic population, the extracted travel times were assigned to individuals. The times thus correspond to a base case scenario in the near future when the bicycle highway is not yet available. The ranges of commute distances and times extracted across modes are indicated in Table 6.2

Table 6.2: Distance and time ranges of commutes in the scenario dataset (n = 450) by mode (extracted from Google Maps)

Mode	Distance (km)	Time (min)
Walk	0.3 – 17.6	3.4 – 220.7
Bicycle	0.3 – 18.0	0.9 – 66.0
Transit	0.3 – 21.3	3.3 – 73.8
Auto	0.3 – 27.1	1.2 – 39.2

Table 6.2 shows that the commutes considered in the dataset are not too long, and hence form an excellent basis for scenario analyses on bicycle commuting. The dataset thus established was then used for the prediction of commute mode shares in scenarios with the bicycle highway.

6.2 Prediction

In this thesis, the scenario dataset described in the previous section was used to estimate the impact of Munich's pilot bicycle highway. This section details the attempted scenarios and the results of the predictions.

As discussed in chapter 5, based on the data available, the only attribute related to transport facilities that could be included in the model was travel time. Consequently, it was the only attribute that could be varied for any predictions related to bicycle highways. Therefore, scenarios with varying travel times calculated based on varying average bicycling speeds were tested.

As described in chapter 2, the bicycle highways proposed in Munich are expected to enable maximum speeds up to 30 km/h (higher for pedelecs and e-bikes). The average speed of bicycle commutes reported in the model dataset, as indicated in Table 5.3, was 13.7 km/h. Whereas, an average bicycling speed of 16.9 km/h was calculated for the commutes of individuals in the scenario dataset extracted from Google Maps. These speeds are from 2008 and 2017 travel times respectively, but both do not account for bicycle highways. Hence, for a future scenario with bicycle highways, the average bicycle speed is bound to increase for various reasons. In this context, the following scenarios were predicted on the scenario dataset by applying the commuter mode choice model:

1. Base case scenario – Travel times as extracted from Google Maps
2. Scenario 18 – Travel times based on an average bicycling speed of 18 km/h
3. Scenario 20 – Travel times based on an average bicycling speed of 20 km/h
4. Scenario 22 – Travel times based on an average bicycling speed of 22 km/h

In the base case scenario, the scenario dataset was used for prediction as it is, i.e. with commute times for trips along the highway corridor as obtained from Google Maps. For the other three scenarios, travel times by bicycle were modified based on assumed average speeds and the commute distances obtained from Google Maps. Travel times by walk, auto and transit were kept unchanged. The change in bicycle commute times computed across the four scenarios can be seen in Table 6.3.

Table 6.3: Bicycle commute times across scenarios

Scenario	**Minimum (min)**	**Mean (min)**	**Maximum (min)**
Base case	0.9	19.7	66.0
Scenario 18	1.0	19.0	60.1
Scenario 20	1.0	17.1	54.1
Scenario 22	0.9	15.5	49.2

With travel time changes as summarized in the above table, choice probabilities were predicted as per the MNL framework described in section 4.1.1 for each scenario by computing utilities using the parameters estimated in section 5.3.2. The resulting mode shares in each of the scenarios are depicted in Fig. 6.4 below.

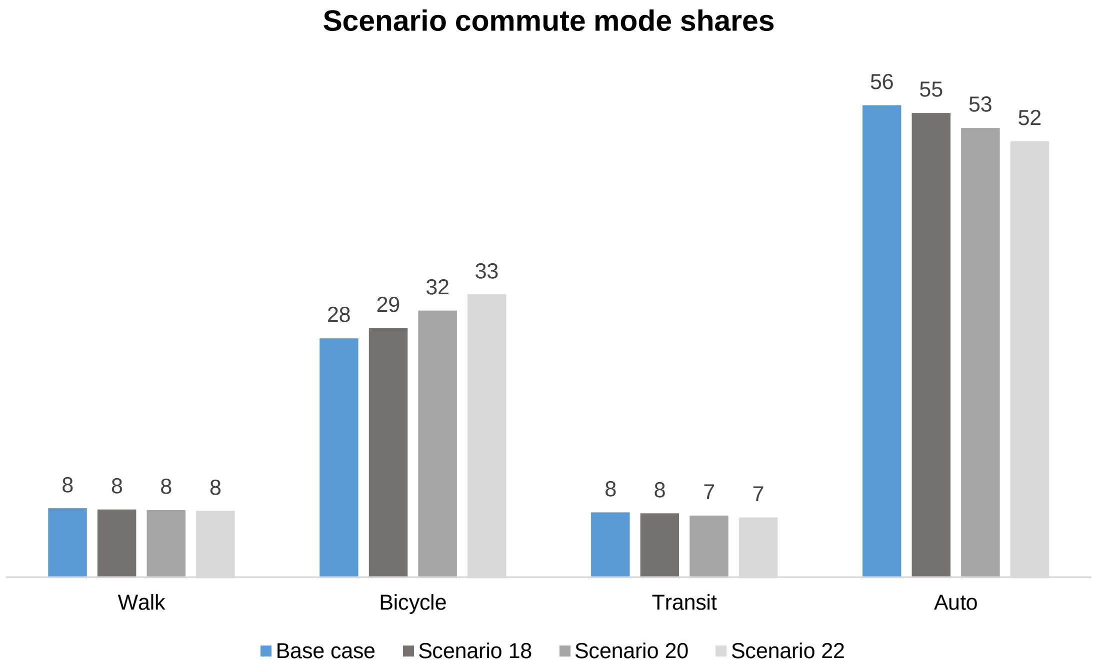

Fig. 6.4: Changes in commute mode shares predicted for scenarios

From the predictions shown in Fig. 6.4, it can be seen that the propensity of bicycle commuting increases with travel-time reductions. The reduction in bicycle travel times seem to most strongly affect the choice of driving to work, while little impact is seen on transit and walk commutes. Although average bicycling speeds of 20 km/h and 22 km/h may be unrealistic, the predictions for these scenarios help in understanding the underlying sensitivities of the model. The shares predicted for the base case scenario are understandably different from the mode shares of urban commutes computed from MiD 2008 data (see Table 5.7) as the scenario dataset was limited to few zones along the bicycle highway corridor in Munich, while the model dataset contained commute trips from all urban regions of the country. However, as one would expect, auto commutes are predicted to have the highest share, and walk commutes the lowest. Interestingly, a higher share of bicycling relative to transit is predicted. This can be explained to be caused by the predominantly short commutes in the scenario dataset as seen in Fig. 6.5.

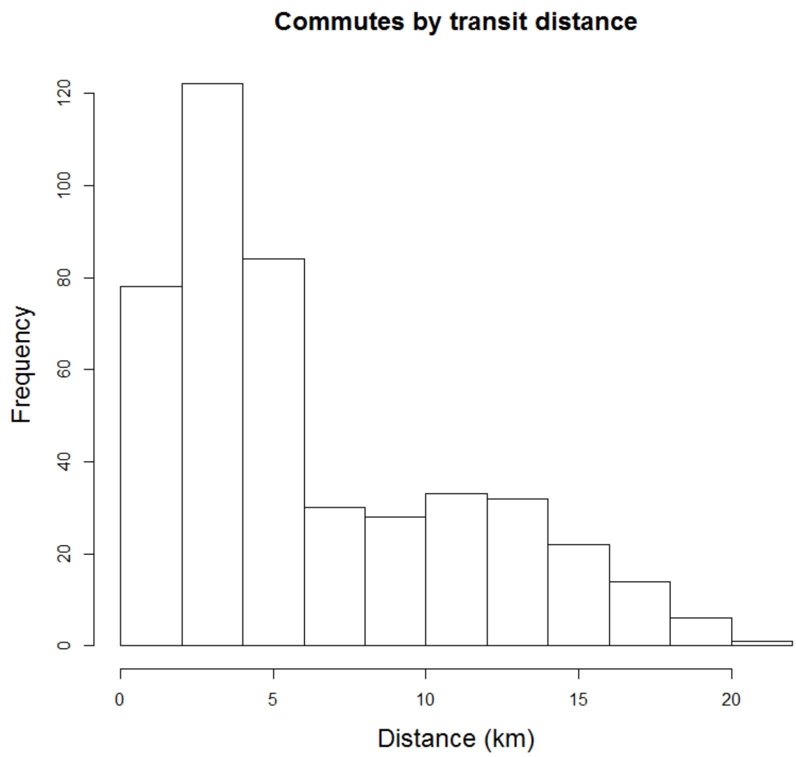

Fig. 6.5: Distribution of scenario dataset commutes by transit distance

Evidently, the majority of the commutes are short, while transit travel is attractive for long commutes (Kuhnimhof et al., 2010). Further, as bicycling becomes more attractive, transit share decreases, however the decline is sharper among auto commutes. This is perhaps because the predicted 8% transit commutes correspond to captive riders, which also explains why the shift from auto travel is stronger.

Moreover, it is also worth noting that the scenarios here only consider the travel time benefit of bicycle highways. While, bicycle highways in reality would have a considerable impact on individuals' attitudes towards bicycling due to increased perceptions of safety, convenience, and comfort, and reduced risks due to the absence of interaction with motorized traffic, attributes that are not captured by the model. The estimated impact can therefore be considered to be conservative. However, the predictions obtained based on the model developed in the thesis forecast an increase in bicycle mode shares and a corresponding decrease in auto shares for an increase in average bicycling speeds due to dedicated cycling infrastructure like bicycle highways.

7 Conclusion

This thesis attempted to assess the impact of a pilot bicycle highway proposed for Munich on the region's commuter mode share. For this, a commuter mode choice model was built for choice among the alternatives – walk, bicycle, transit and auto. The model was based on urban commute data from a national household travel survey conducted in 2008, Mobilität in Deutschland (MiD 2008). Constrained by the information available, the following attributes were used to specify the model – age, gender, income, household size, number of children (< 18 years) in household, auto availability, bicycle availability and travel time.

The model was estimated with a logit modeling framework using R's *mlogit* package. Both MNL and NL estimations were performed, however, the data lacked the required level of detail to obtain reliable results for the NL estimation. Nevertheless, the MNL model was successfully estimated. The estimation yielded statistically significant and theoretically consistent results for all the variables except *income* and *number of children*. The variable *number of children* was further found to be correlated with *household size*. The effect of *age* on walking and that of *bicycle availability* on transit were also statistically insignificant. To ensure a minimum 95% level of significance, these variables were removed from the model. When tested for accuracy with the same dataset, the mode shares predicted using the estimated parameters were within $\pm$2% of the dataset's actual shares.

The estimated model was applied to scenarios by varying the effects of the bicycle highway on bicycling times. The dataset considered for this purpose was prepared with the help of information provided by the Modeling Spatial Mobility research group. The group generated a synthetic population for the Munich metropolitan region and assigned the population to geographical zones for their research on developing an integrated land-use and transport model. Information about individuals assigned to be living and working in zones that lie along the bicycle highways corridor was obtained for this thesis' scenario analysis as they would have the option to use it when available. Most of the required socio-economic variables were available from this dataset. The missing variables of bicycle availability and auto availability were engineered through regressions using the available variables. Further, travel times and distances were extracted from Google Maps through a Java script developed by the research group for a similar purpose. An average bicycling speed of 16.9 km/h was calculated for the extracted commute times. Considering this as the base case, three scenario cases were designed – with average bicycling speeds of 18 km/h, 20 km/h and 22 km/h respectively. Choice probabilities were predicted for each scenario by applying the estimated model parameters. The resulting differences in mode shares showed an increase in bicycling share across scenarios with a corresponding opposite effect on auto shares. Little impact was seen on transit and walk shares. Further, across all scenarios, auto shares were the highest, followed by bicycle. Transit and walk had equally low shares of around 8%. The low transit shares could be explained as the effect of the predominantly short length of commutes in the scenario dataset.

The model's predictions suggest that reduction in bicycling times due to dedicated infrastructure like bicycle highways increases the propensity of bicycle commuting, and reduces the relative attractiveness of auto travel. However, caution should be exercised while interpreting the results as the model has neither been calibrated, nor validated before its application to the scenarios within the scope of this thesis. Nevertheless, the model does not consider the impacts of bicycle highways on attributes like perceptions of increased safety, comfort, convenience and reduced risks which would further improve bicycling utility, indicating that the prediction is perhaps on the conservative side.

Limitations

Some of the limitations of the work carried out in this thesis are listed below:

Model limitations

- The model was based on a dataset containing commute trips from all urban areas in Germany. Availability of such information exclusive to the Munich region would perhaps have resulted in a more adequate model for the scenarios attempted. However, as the model could not be validated within the scope of this thesis, it is difficult to comment on the adequacy of the current model.
- Due to the lack of a more recent data source, the model was built on data collected in 2008 assuming no changes in mobility behavior since then. But the last decade has seen the development of many alternative mobility options like shared-use mobility, which would have caused some changes in people's travel behavior.
- The unit of analysis for the model was an individual commute trip, implying that multimodal travel behavior could not be considered as would have been possible when considering tours and trip chains.
- A common choice set was assumed for all individuals, while in reality certain groups would not have, or consider certain modes.
- Travel times were available only in the form of trip duration (start to end). While, mode choice is generally more sensitive to its components like in-vehicle time, access time and egress time.
- Travel times for modes that were not chosen were engineered based on the average speeds reported. They should ideally be obtained by performing a trip assignment exercise between the origins and destinations.

- The effect of travel costs could not be considered due to unavailability of cost information. As a result, the model cannot be tested for scenarios involving changes in costs like fares, parking charges, tolls, etc.
- Many attributes of transit choice could not be included either due to lack of information (e.g. number of transfers, waiting time, access and egress times, etc.) or to preserve the sample size (information on transit subscriptions and distance from households to nearest transit stop were missing for 30-40% of the dataset).
- The model only involves objective variables related to the individuals and the mode of transport (travel time). While, the choice to bicycle is perhaps better explained by subjective variables related to attitudes, perceptions and habits.
- Due to the lack of adequately fine data, estimations considering correlations between alternatives could not be performed.

Scenario limitations

- The model was applied on the scenarios without prior calibration or validation. This makes it difficult to comment on the validity of the predictions.
- Due to the lack of a complete dataset, some variables were approximated through regressions whose validity cannot be tested.
- Field measurements of average speeds attainable on bicycle highways would allow for exact estimation of the potential impacts.

In future research, an attempt should be made to address the above limitations in order to improve the reliability and validity of the results predicted in similar studies.

List of References

20minuten. (2014). Mehrspurige Velobahn für Winterthur. Retrieved from http://www.20min.ch/schweiz/zuerich/story/17468626?redirect=mobi&nocache=0.3520287969149649

ADFC München e.V. (2014). *Radschnellweg: München Nord - Hauptbahnhof*. Retrieved from https://www.adfc-muenchen.de/fileadmin/user_upload/kv-muenchen/radverkehr/2014/Radschnellweg/Muenchen/ADFC-Muenchen_Radschnellweg_Nord.pdf

Banister, C. (1990). Planning for cycling. *Planning Practice and Research*, *5*(1), 27–29. doi:10.1080/02697459008722763

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. Massachusetts: MIT Press.

Bergström, A., & Magnusson, R. (2003). Potential of transferring car trips to bicycle during winter. *Transportation Research Part A*, *37*(8), 649–666. Retrieved from http://www.sciencedirect.com/science/article/pii/S0965856403000120

Bicycling.com. (2015). Germany Opens Bicycle-Only Autobahn. Retrieved from http://www.bicycling.com/culture/advocacy/germany-opens-bicycle-only-autobahn

Bowman, J. L. (2014). *Incorporating Bicycling into Activity-based Regional Travel Forecasting Models in Denmark: Identified Needs and Proposed Solutions*. Draft 3. Retrieved from Danish Road Directorate (Vejdirektoratet) website: http://jbowman.net/papers/Bicycle%20Modeling%20v3%20Letter.pdf

Broach, J. P. (2016). *Travel Mode Choice Framework Incorporating Realistic Bike and Walk Routes* (Dissertation). Portland State University, Portland, Oregan. Retrieved from http://pdxscholar.library.pdx.edu/open_access_etds/2702/

Buehler, R. (2012). Determinants of bicycle commuting in the Washington, DC region: The role of bicycle parking, cyclist showers, and free car parking at work. *Transportation Research Part D*, *17*(7), 525–531. Retrieved from http://www.sciencedirect.com/science/article/pii/S1361920912000594

Buekers, J., Dons, E., Elen, B., & Panis, L. I. (2015). Health impact model for modal shift from car use to cycling or walking in Flanders: application to two bicycle highways. *Journal of Transport & Health*, *2*(4), 549–562. Retrieved from http://www.sciencedirect.com/science/article/pii/S221414051500674X

Bundesministerium für Verkehr und digitale Infrastruktur. (n.d.[2009]). Mobilität in Deutschland: Publikationen zur Erhebungswelle 2008. Retrieved from http://www.mobilitaet-in-deutschland.de/mid2008-publikationen.html

Bundesministerium für Verkehr und digitale Infrastruktur. (2016). Voraussetzungen für die Fahrerlaubnis und Geltungsdauer.

City of Munich, Department of Urban Planning and Building Regulation. (2006). *Transport Development Plan: Information about urban development.*

Croissant, Y. (n.d.[2011]). *Estimation of multinomial logit models in R: The mlogit Packages*. Retrieved from Universit´e de la R´eunion website: https://cran.r-project.org/web/packages/mlogit/vignettes/mlogit.pdf

Cycle Super Highways, Capital Region Copenhagen. (n.d.). Download press photos. Retrieved from http://supercykelstier.dk/pressphotos/

Dill, J., & Wardell, E. (2007). Factors Affecting Worksite Mode Choice: Findings from Portland, Oregon. *Transportation Research Record: Journal of the Transportation Research Board*, *1994*, 51–57. Retrieved from http://trrjournalonline.trb.org/doi/abs/10.3141/1994-07

Dunlap, M. (2015). *How The Weather, Land Use, and Infrastructure Influence Non-Motorized Mode Choices* (Master's Dissertation). University of Washington, Waschington, D.C. Retrieved from https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/33121/Dunlap_washington_0250O_14306.pdf;sequence=1

Engbers, L. H., & Hendriksen, I. J. (2010). Characteristics of a population of commuter cyclists in the Netherlands: perceived barriers and facilitators in the personal, social and physical environment. *International Journal of Behavioral Nutrition and Physical Activity*, *7*(89), 1–5. Retrieved from http://ijbnpa.biomedcentral.com/articles/10.1186/1479-5868-7-89

European Cyclists' Federation. (n.d.[2014]). Fast Cycling Routes: Towards fast commuting. Retrieved from https://ecf.com/what-we-do/urban-mobility/fast-cycling-routes

European Cyclists' Federation. (2014). *Fast Cycling Routes: towards barrier-free commuting*. Factsheet. Brussels, Belgium.

Fernández-Heredia, Á., Jara-Díaz, S., & Monzón, A. (2016). Modelling bicycle use intention: the role of perceptions. *Transportation*, *43*(1), 1–23. Retrieved from http://link.springer.com/article/10.1007/s11116-014-9559-9

Follmer, R., Gruschwitz, D., Jesske, B., Quandt, S., Lenz, B., Nobis, C.,. . . Mehlin, M. (2010). *Mobilität in Deutschland 2008: Ergebnisbericht. Struktur - Aufkommen - Emissionen - Trends*. Bonn, Berlin.

German Aerospace Center (DLR). (2012). Mobility in Germany 2008. Retrieved from http://daten.clearingstelle-verkehr.de/223/

Goudappel Coffeng. (n.d.a). Fietssnelweg - een definitie. Retrieved from http://www.fietssnelwegen.nl/index.php?m=1

Goudappel Coffeng. (n.d.b). Overzicht fietssnelwegen in Nederland. Retrieved from http://www.fietssnelwegen.nl/index.php?m=2

Halldórsdóttir, K., Christensen, L., Jensen, T. C., & Prato, C. G. (2011). Modelling Mode Choice in Short Trips: Shifting from Car to Bicycle. In *European Transport Conference 2011: Seminars.* Retrieved from https://trid.trb.org/view.aspx?id=1258637

Hamre, A., & Buehler, R. (2014). Commuter Mode Choice and Free Car Parking, Public Transportation Benefits, Showers/Lockers, and Bike Parking at Work: Evidence from the Washington, DC Region. *Journal of Public Transportation*, *17*(2), 67–91. doi:10.5038/2375-0901.17.2.4

Heinen, E., Maat, K., & van Wee, B. (2011). The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances. *Transportation Research Part D: Transport and Environment*, *16*(2), 102–109. doi:10.1016/j.trd.2010.08.010

Heinen, E., Maat, K., & van Wee, B. (2013). The effect of work-related factors on the bicycle commute mode choice in the Netherlands. *Transportation*, *40*(1), 23–43. Retrieved from http://link.springer.com/article/10.1007/s11116-012-9399-4

Heinen, E., van Wee, B., & Maat, K. (2010). Commuting by Bicycle: An Overview of the Literature. *Transport Reviews: A Transnational Transdisciplinary Journal*, *30*(1), 59–96. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/01441640903187001

Hensher, D. A., & Bullock, R. G. (1979). Price elasticity of commuter mode choice: Effect of a 20 per cent rail fare reduction. *Transportation Research Part A*, *13*(3), 193–202. Retrieved from http://www.sciencedirect.com/science/article/pii/0191260779900712

Hensher, D. A., & Rose, J. M. (2007). 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*, *41*(5), 428–443. Retrieved from http://www.sciencedirect.com/science/article/pii/S0965856406001145

Hensher, D. A., Rose, J. M., & Greene, W. H. (2005). *Applied Choice Analysis: A Primer*. New York, USA: Cambridge University Press.

Hess, D. B. (2001). Effect of free parking on commuter mode choice: Evidence from travel diary data. *Transportation Research Record*, *1753*, 35–42. Retrieved from http://trrjournalonline.trb.org/doi/abs/10.3141/1753-05

Hunt, J., & Abraham, J. (2007). Influences on bicycle use. *Transportation*, *34*(4), 453–470. Retrieved from http://link.springer.com/article/10.1007/s11116-006-9109-1

infas Institut für angewandte. (n.d.). Mobilität in Deutschland. Retrieved from http://www.mobilitaet-in-deutschland.de/index.html

Koppelman, F. S., & Bhat, C. (2006). *A Self Instructing Course in Mode Choice Modeling: Multinomial and Nested Logit Models*. Retrieved from U.S. Department of Transportation, Federal Transit Administration website: http://www.caee.utexas.edu/prof/bhat/COURSES/LM_Draft_060131Final-060630.pdf

Kuhnimhof, T., Chlond, B., & Huang, P.-C. (2010). Multimodal Travel Choices of Bicyclists: Multday Data Analysis of Bicycle Use in Germany. *Transportation Research Record: Journal of the Transportation Research Board*, *2190*, 19–27. Retrieved from http://trrjournalonline.trb.org/doi/abs/10.3141/2190-03

Landeshauptstadt München. (n.d. [2016]). Warum Radlhauptstadt? Retrieved from https://radlhauptstadt.muenchen.de/ueber-uns/

Landeshauptstadt München, Referat für Stadtplanung und Bauordnung. (2010). *Mobilität in Deutschland (MiD): Alltagsverkehr in München, im Münchner Umland und im MVV-Verbundraum*. München.

London Cycling Campaign. (n.d.). Cycle superhighways. Retrieved from http://lcc.org.uk/pages/cycle-superhighways

MacKinder, I. H., Evans, S. E., & May, R. (1975). The distribution of household income in trip generation. *Traffic Engineering & Control*, *16*(12). Retrieved from https://trid.trb.org/view.aspx?id=46552

Maldonado-Hinarejos, R., Sivakumar, A., & Polak, J. W. (2014). Exploring the role of individual attitudes and perceptions in predicting the demand for cycling: a hybrid choice modelling approach. *Transportation*, *41*(6), 1287–1304. Retrieved from http://link.springer.com/article/10.1007/s11116-014-9551-4

Moore, R. (2016). Superhighway to cycling heaven – or just a hell of a mess? Retrieved from https://www.theguardian.com/lifeandstyle/2016/aug/07/london-cycle-superhighways-heaven-or-hell

Mühlfenzl, M. (2016, April 1). Auf dem Radl-Highway in die Arbeit. *Sueddeutsche Zeitung*. Retrieved from http://www.sueddeutsche.de/muenchen/landkreismuenchen/radschnellwege-auf-dem-radl-highway-in-die-arbeit-1.2929326

Munizaga, M. A., & Ortúzar, J. d. D. (1999). Nested Logit Modelling: Some Hard Facts. *27th European Transport Conference*, *P324*, 25–36.

Muñoz, B., Monzon, A., & Daziano, R. A. (2016). The Increasing Role of Latent Variables in Modelling Bicycle Mode Choice. *Transport Reviews*, *36*(6), 737–771. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/01441647.2016.1162874

Nankervis, M. (1999). The effect of weather and climate on bicycle commuting. *Transportation Research Part A*, *33*(6), 417–431. Retrieved from http://www.sciencedirect.com/science/article/pii/S0965856498000226

Noland, R. B., & Kunreuther, H. (1995). Short-run and long-run policies for increasing bicycle transportation for daily commuter trips. *Transport Policy*, *2*(1). Retrieved from http://www.sciencedirect.com/science/article/pii/0967070X9593248W

Ortúzar, J. d. D., Iacobelli, A., & Valeze, C. (2000). Estimating demand for a cycle-way network. *Transportation Research Part A*, *34*(5), 353–373. Retrieved from http://www.sciencedirect.com/science/article/pii/S0965856499000403

Ortúzar, J. d. D., & Willumsen, L. G. (2011). *Modelling Transport* (Fourth Edition). Chichester, West Sussex, United Kingdom: John Wiley & Sons, Ltd.

O'Sullivan, F. (2016). Germany Launches Its National 'Bike Autobahn' Cycle Network. Retrieved from http://www.citylab.com/commute/2016/01/germany-launches-its-national-bike-autobahn-cycle-network/422451/

Parkin, J., Wardman, M., & Page, M. (2008). Estimation of the determinants of bicycle mode share for the journey to work using census data. *Transportation*, *35*(1), 93–109. Retrieved from http://link.springer.com/article/10.1007/s11116-007-9137-5

Paul, F., Bogenberger, K., & Fink, B. (2016). Evaluation of Munich’s Cycle Route Planner: Data Analysis and Customer Survey. In *International Scientific Conference on Mobility and Transport Transforming Urban Mobility.*

Pfertner, M. (2015). *Radschnellwege zur Förderung nachhaltiger Mobilität im Raum München* (Bachelor's Thesis). Technische Unviersität München, Munich. Retrieved from https://mediatum.ub.tum.de/doc/1304073/1304073.pdf

Planungsverband Äußerer Wirtschaftsraum München. (2015a). *Radschnellverbindungen in München und Umland*. Retrieved from http://www.pv-muenchen.de/index.php?id=1515,312&suche=

Planungsverband Äußerer Wirtschaftsraum München. (2015b). *Pressemitteilung: 14 Radschnellwege in München und Umland*. München. Retrieved from http://www.pv-muenchen.de/index.php?id=0,225

pressedienst-fahrrad GmbH. (2014). Themenblatt: Die Fahrradwelt in Zahlen. Retrieved from http://www.pd-f.de/wp-content/uploads/kalins-pdf/singles/themenblatt-die-fahrradwelt-in-zahlen.pdf

Regionalverband Ruhr. (n.d.). metropoleruhr: RS1. Der schnellste weg durchs Revier. Retrieved from https://www.rs1.geoportal.ruhr/

Rodríguez, D. A., & Joo, J. (2004). The relationship between non-motorized mode choice and the local physical environment. *Transportation Research Part D*, *9*(2), 151–173. Retrieved from http://www.sciencedirect.com/science/article/pii/S1361920903000889

Ruhr Regional Association (Regionalverband Ruhr). (n.d.[2016]). *Feasibility Study: Radschnellweg RS1*. Abstract. Retrieved from http://www.rs1.ruhr/fileadmin//user_upload/RS1/pdf/RS1_abstract_web_21.11.2016.pdf

Schwägerl, C. (2016). Beyond the Autobahn: Germany’s new bike highways. Retrieved from https://www.greenbiz.com/article/beyond-autobahn-germanys-new-bike-highways

Schwanen, T., & Mokhtarian, P. L. (2005). What affects commute mode choice: neighborhood physical structure or preferences toward neighborhoods? *Journal of Transport Geography*, *13*(1), 83–99. Retrieved from http://www.sciencedirect.com/science/article/pii/S0966692304000894

Sekretariatet for Supercykelstier. (n.d.a). Supercykelstier: The concept. Retrieved from http://supercykelstier.dk/the-concept/

Sekretariatet for Supercykelstier. (n.d.b). Supercykelstier: The project. Retrieved from http://supercykelstier.dk/the-project/

Sekretariatet for Supercykelstier. (n.d.c). Supercykelstier: The routes. Retrieved from http://supercykelstier.dk/the-routes/

Sekretariatet for Supercykelstier. (2014). *FARUMRUTEN: EVALUERING*. Retrieved from http://supercykelstier.dk/wp-content/uploads/2016/03/Evaluering-af-Farumruten.pdf

Sener, I. N., Eluru, N., & Bhat, C. R. (2009). Who Are Bicyclists? Why and How Much Are They Bicycling? *Transportation Research Record: Journal of the Transportation Research Board*, *2134*, 63–72. Retrieved from http://trrjournalonline.trb.org/toc/trr/2134

Sigurdardottir, S. B., Kaplan, S., Møller, M., & Teasdale, T. W. (2013). Understanding adolescents' intentions to commute by car or bicycle as adults. *Transportation Research Part D*, *24*, 1–9. Retrieved from http://www.sciencedirect.com/science/article/pii/S1361920913000710

SRF Schweizer Radio und Fernsehen. (2015). Velobahn in St. Gallen: Was kostet das denn? Retrieved from http://www.srf.ch/news/regional/ostschweiz/velobahn-in-st-gallen-was-kostet-das-denn

Technical University Munich. (2016). Simple Integrated Land-Use Orchestrator: Creating a Synthetic Population with German Data. Retrieved from http://silo.zone/synPopDE.html

Thiemann-Linden, J., & Boeckhout, S. Van. (2012). *Cycle Highways* (Cycling Expertise – Infrastructure No. I-12).

Tilahun, N. Y., Levinson, D. M., & Krizek, K. J. (2007). Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. *Transportation Research Part A*, *41*(4), 287–301. Retrieved from http://www.sciencedirect.com/science/article/pii/S096585640600108X

Transport for London. (2011). *Barclays Cycle Superhighways: Evaluation of Pilot Routes 3 and 7*. Retrieved from https://www.whatdotheyknow.com/request/162841/response/402026/attach/3/BCS%20pilot%20evaluation%20report.pdf

Transport for London. (n.d.[2016]). Cycle Superhighways. Retrieved from https://tfl.gov.uk/modes/cycling/routes-and-maps/cycle-superhighways

Tscharnke, K. (2015). Kommt jetzt die Radl-Autobahn für München?: Planungen für Radschnellwege. Retrieved from http://www.tz.de/muenchen/stadt/muenchen-ort29098/muenchen-planungen-radschnellwege-5258047.html

Twaddle, H. A. (2011). *Stated preference survey design and pre-test for valuing influencing factors for bicycle use* (Master's thesis). Technische Unviersität München, München. Retrieved from https://mediatum.ub.tum.de/doc/1130707/1130707.pdf van der Zee, R. (2016). Could intercity cycle highways revolutionise the daily commute? Retrieved from https://www.theguardian.com/cities/2016/jun/30/intercity-cycle-highways-revolutionise-daily-commute

Viton, P. A. (2015). *Discrete-Choice Logit Models with R*. Retrieved from http://facweb.knowlton.ohio-state.edu/pviton/courses2/crp5700/5700-mlogit.pdf

Walter, D. (2015). 17 Kilometer Rad-Schnellweg geplant. Retrieved from http://www.merkur.de/bayern/studie-landkreise-stadt-muenchen-kilometer-rad-schnellweg-geplant-5246469.html

Wardman, M., Hatfield, R., & Page, M. (1997). The UK national cycling strategy: can improved facilities meet the targets? *Transport Policy*, *4*(2), 123–133. Retrieved from http://www.sciencedirect.com/science/article/pii/S0967070X97000115

Wardman, M., Tight, M., & Page, M. (2007). Factors influencing the propensity to cycle to work. *Transportation Research Part A*, *41*(4), 339–350. Retrieved from http://www.sciencedirect.com/science/article/pii/S0965856406001212

Washbrook, K., Haider, W., & Jaccard, M. (2006). Estimating commuter mode choice: A discrete choice analysis of the impact of road pricing and parking charges. *Transportation*, *33*(6), 621–639. Retrieved from http://link.springer.com/article/10.1007/s11116-005-5711-x

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2011). *Statistical and Econometrical Methods for Transportation Data Analysis* (2nd ed.). New York, USA: CRC Press.

Willis, D. P., Manaugh, K., & El-Geneidy, A. (2015). Cycling under influence: Summarizing the influence of attitudes, habits, social environments and perceptions on cycling for transportation. *International Journal of Sustainable Transportation*, *9*(8), 565–579.

Wootton, H. J., & Pick, G. W. (1967). A model for trips generated by households. *Journal of Transport Economics and Policy*, 137–153.

List of Abbreviations

RP	Revealed Preference
SP	Stated Preference
MiD	Mobilität in Deutschland
MNL	Multinomial Logit
NL	Nested Logit

List of Figures

List of Tables

Appendix A: Explanatory variables used in similar studies

Study		Noland & Kunreuther (1995): Short-run and long-run policies for increasing bicycle transportation for daily commuter trips	Wardman et.al. (1997): The UK national cycling strategy: can improved facilities meet the targets?	Rodríguez and Joo (2004): The relationship between non-motorized mode choice and the local physical environment	Wardman et.al. (2007): Factors influencing the propensity to cycle to work	Hamre & Buehler (2014): Commuter Mode Choice and Free Car Parking, Public Transportation Benefits, Showers/Lockers, and Bike Parking at Work: Evidence from the Washington, DC Region
Country		USA	UK	USA	UK	USA
Data type		RP	SP	RP	Combined RP and SP	RP
Model structure		MNL	MNL	NL	MNL	MNL
Alternatives		Walk Bike Auto Transit	Car Bus Bike	No-vehicle: Walk Collective modes: Carpool, bus, park-ride Individual modes: Bicycle, auto	Car driver Car passenger Bus Walk Bike	Walk Bike Transit Drive alone
Explanatory variables	Characteristics of trip maker	Gender, Income, Number of motor vehicles owned, Bicycling competency, Hours of exercise per week.	-	Age, Gender, Student (y/n), Number of household vehicles, Driving license posession(y/n), Gross population density of the residing block (persons/sq.km), Altitude at house location(m above sea level).	Age, Gender, Income, Type of work (skilled /unskilled /clerical), Financial incentives for cycling to work, Proportion of general population that cycled, Proportion of colleagues cycling, Availability of a company car, Use of car for work.	Age, Gender, Income, Number of children in household, Number of cars per household member, Number of bicycles per household member, Home zone population density.
	Characteristics of journey	-	Weather, Shower facilities at destination.	-	Shower facilities destination.	Commute distance, Season, Showers, locker facilities at destination.
	Characteristics of transport facility	Travel time, Bicycle parking availability, Perceived cost, convenience, risk, and comfort.	Travel time, Travel cost, Secure bicycle parking facilities at destination.	Travel time Number of vehicles during peak hour Walking and bicycling paths % of shortest route with side-walk to bus-stop for transit and campus for walk Slope	Travel time, Costs, Parking availability, Ratings of cycle trips in terms of hilliness, air pollution, noise, danger from traffic, personal security, tiredness and cycling ability.	Parking, vanpooling facilities at destination, Number of metro stations in home zone, Centerline miles of bike lanes per 1000 residents in home zone.

Appendix B: Dataset arranged in wide format (sample)

mode	age	gender	income	hh_size	n_child	auto_avail	bike_avail	peak_hr	ttime.walk	ttime.bike	ttime.auto	ttime.transit	weight
auto	37	0	745	3	1	0.5	0	0	35.4	12.52	15	8.19	1.212781412
auto	44	0	979	4	1	0.67	1	1	70.81	25.04	15	16.37	1.225885756
bike	18	1	944	4	1	0.67	1	1	12.17	10	1.53	2.81	1.056362832
walk	25	1	961	4	0	0.67	0	0	3	0.88	0.31	0.57	0.793334761
walk	59	0	938	4	0	0.67	0	0	3	0.88	0.31	0.57	0.296812522
walk	19	0	964	4	0	0.67	0	0	2	0.88	0.31	0.57	0.561848364
auto	30	0	506	2	0	0.5	1	0	212.42	75.11	30	49.11	1.598395679
auto	36	0	2248	3	1	1	1	1	35.4	12.52	10	8.19	1.589899243
auto	41	1	2980	4	2	0.5	1	1	31.93	11.29	7	7.38	0.998793062
transit	51	1	1156	3	0	0.33	1	0	291.55	103.09	36.53	60	1.288087619
bike	52	0	1009	3	0	0.33	1	1	133.91	35	16.78	30.96	1.30958267
auto	52	1	2883	2	0	0.5	1	0	82.61	29.21	10	19.1	0.509690759
transit	42	0	1413	2	1	0	1	1	280.25	99.09	35.11	58	1.727820574
auto	42	0	1485	5	3	1	1	1	147.58	52.18	20	34.12	0.646460642
auto	48	0	2206	5	2	1	1	1	212.42	75.11	30	49.11	0.572866566
transit	41	1	2222	3	1	0.5	1	0	728.7	257.66	91.3	109	1.144310247
bike	61	0	2686	2	0	1	1	0	48.7	20	6.1	11.26	0.399752224
auto	49	1	2313	2	0	1	1	0	295.03	104.32	30	68.21	1.066741174
auto	38	0	2295	1	0	1	0	1	106.21	37.55	30	24.56	3.160936582
auto	22	0	4994	3	0	0.67	1	1	47.2	16.69	10	10.91	1.050131525
auto	52	1	2270	5	3	0.5	0	1	200.62	70.94	15	46.39	0.449954966
transit	58	1	946	4	0	0.5	1	1	168.2	59.47	21.07	60	0.029813013
transit	56	0	980	4	0	0.5	1	1	168.2	59.47	21.07	60	0.036135898
transit	29	0	964	4	0	0.5	1	0	112.17	39.66	14.05	60	0.176373788
transit	61	1	2107	2	0	1	0	1	1008.94	356.75	126.41	91	0.61271892
auto	45	1	3057	4	2	1	1	1	236.02	83.46	30	54.57	0.476823487
auto	54	0	2211	2	0	1	0	1	8.32	2.94	5	1.92	1.359310671
auto	50	1	1084	4	0	0.75	1	0	330.43	116.84	45	76.4	1.211674797
auto	49	1	4994	5	1	1.25	1	0	283.23	100.15	25	65.49	0.243602402
auto	44	1	703	3	0	0.67	0	0	354.04	125.18	35	81.86	1.000001172
auto	21	1	1770	3	0	1	0	0	283.23	100.15	29	65.49	0.901400921
auto	42	0	2137	4	2	1	1	0	531.06	187.77	60	122.79	0.632041443
auto	44	1	2113	4	2	1	1	1	413.04	146.05	40	95.5	0.621667327
auto	25	1	1339	2	0	1	0	1	59.01	20.86	10	13.64	1.74871226
auto	29	0	1402	2	0	1	0	0	177.02	62.59	20	40.93	1.686213126
auto	33	1	2026	2	0	1	1	1	106.21	37.55	15	24.56	1.91049845
auto	40	1	1786	2	1	1	1	1	123.98	43.84	15	28.66	1.328827596
auto	52	1	1891	1	0	1	0	0	177.02	62.59	27	40.93	1.292458564
auto	48	0	771	3	0	0.67	1	0	236.02	83.46	20	54.57	0.912871334
auto	40	0	1616	4	2	1	1	1	271.43	95.97	30	62.76	0.857947881
auto	42	1	1599	4	2	1	1	1	129.81	45.9	20	30.01	0.907055232
auto	30	0	1701	1	0	1	1	1	413.04	146.05	30	95.5	9.145156016
bike	28	1	2558	1	0	1	1	1	18.26	5	2.29	4.22	4.033455426
transit	16	0	1489	4	2	0.5	1	1	157.02	55.52	19.67	42	0.646755547
auto	40	0	1405	4	2	0.5	1	1	118.01	41.73	15	27.29	0.758147717
auto	41	1	1221	4	2	1	1	1	94.41	33.38	25	21.83	0.688251671
auto	53	0	1872	5	2	0.33	1	1	59.01	20.86	15	13.64	0.30593719
auto	23	0	935	4	0	1	1	0	542.86	191.95	80	125.51	1.435941582
auto	51	0	2973	2	0	1	1	0	50.81	17.96	9	11.75	0.904518642

Appendix C: Result output of multinomial logit estimation of the model

```
Call:
mlogit(formula = mode ~ I(bike * bike_avail) + I(walk * bike_avail) +
    I(bike * gender) + I(transit * gender) + I(walk * ttime) +
    I(bike * ttime) + I(auto * ttime) + I(transit * (ttime^2)) |
    age + hh_size + auto_avail | 0, data = mnl_data, weights = mnl_data$weight,
    method = "nr", print.level = 0)

Frequencies of alternatives:
    auto     bike  transit     walk
0.703148 0.101720 0.138444 0.056689

nr method
208 iterations, 0h:0m:6s
g'(-H)^-1g = 1.38E-06
successive function values within tolerance limits

Coefficients :
                          Estimate  Std. Error  t-value  Pr(>|t|)
bike:(intercept)        1.2523e+00  3.2904e-01   3.8061 0.0001412 ***
transit:(intercept)     2.2955e+00  2.2902e-01  10.0231 < 2.2e-16 ***
walk:(intercept)        5.6753e+00  4.6900e-01  12.1010 < 2.2e-16 ***
I(bike * bike_avail)    1.9764e+00  1.9116e-01  10.3387 < 2.2e-16 ***
I(walk * bike_avail)   -4.2510e-01  1.7502e-01  -2.4288 0.0151474 *
I(bike * gender)        2.4380e-01  9.2386e-02   2.6390 0.0083158 **
I(transit * gender)    -3.2741e-01  7.9871e-02  -4.0992 4.146e-05 ***
I(walk * ttime)        -1.5228e-01  7.9581e-03 -19.1357 < 2.2e-16 ***
I(bike * ttime)        -7.4850e-02  3.4939e-03 -21.4233 < 2.2e-16 ***
I(auto * ttime)        -1.8567e-02  2.6062e-03  -7.1242 1.047e-12 ***
I(transit * (ttime^2)) -6.3995e-05  1.1286e-05  -5.6704 1.425e-08 ***
bike:age               -1.3242e-02  4.1860e-03  -3.1635 0.0015591 **
transit:age            -2.0833e-02  3.3034e-03  -6.3066 2.853e-10 ***
walk:age               -1.3448e-02  6.3291e-03  -2.1248 0.0336053 *
bike:hh_size           -8.5931e-02  4.1115e-02  -2.0900 0.0366161 *
transit:hh_size        -2.4153e-01  3.5721e-02  -6.7616 1.365e-11 ***
walk:hh_size           -3.4830e-01  6.3708e-02  -5.4671 4.574e-08 ***
bike:auto_avail        -3.1013e+00  1.6021e-01 -19.3575 < 2.2e-16 ***
transit:auto_avail     -3.4413e+00  1.3448e-01 -25.5890 < 2.2e-16 ***
walk:auto_avail        -2.2565e+00  2.2901e-01  -9.8535 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Log-Likelihood: -4113.2
McFadden R^2:  0.34603
Likelihood ratio test : chisq = 4352.7 (p.value = < 2.22e-16)
```

Appendix D: Result output of logistic regression for *bicycle availability*

```
Call:
glm(formula = bike_avail ~ age + hh_size + n_child + log(income),
    family = quasibinomial, data = train_data_new, weights = train_data_new$weight)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-5.1376   0.2933   0.4958   0.7005   3.3851

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.778299   0.551972  -5.033 4.99e-07 ***
age          0.020896   0.002965   7.047 2.07e-12 ***
hh_size     -0.156819   0.037636  -4.167 3.14e-05 ***
n_child      0.329592   0.062493   5.274 1.39e-07 ***
log(income)  0.471870   0.073777   6.396 1.74e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for quasibinomial family taken to be 1.067299)

    Null deviance: 5585.3  on 4999  degrees of freedom
Residual deviance: 5423.2  on 4995  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 4
```

Appendix E: Result output of multiple linear regression for *auto availability*

```
Call:
lm(formula = auto_avail ~ gender + hh_size + n_child + log(income),
    data = train_data_new, weights = train_data_new$weight)

Weighted Residuals:
    Min      1Q  Median      3Q     Max
-2.5713 -0.2003  0.0746  0.2030  9.2433

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.020157   0.088672   0.227   0.8202
gender       0.016994   0.011208   1.516   0.1295
hh_size     -0.040292   0.006308  -6.387 1.84e-10 ***
n_child      0.021257   0.009656   2.202   0.0277 *
log(income)  0.114601   0.011681   9.811  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.401 on 4995 degrees of freedom
Multiple R-squared:  0.03099,   Adjusted R-squared:  0.03021
F-statistic: 39.93 on 4 and 4995 DF,  p-value: < 2.2e-16
```

Declaration concerning the Master’s Thesis

I hereby confirm that the presented thesis work has been done independently and using only the sources and resources as are listed. This thesis has not previously been submitted elsewhere for purposes of assessment.

Munich, January 24th, 2017

R. Hema Sharanya

Hema Sharanya Rayaprolu