

2011
2012

BEDRIJFSECONOMISCHE WETENSCHAPPEN
master in de verkeerskunde: mobiliteitsmanagement
(Interfacultaire opleiding)

Masterproef

Modelling Public Transport Potential in Relation to Area Characteristics

Promotor :
Prof. dr. Gerhard WETS

Copromotor :
De heer Petrus VAN DER WAERDEN

Wannes Poesen

*Masterproef voorgedragen tot het bekomen van de graad van master in de verkeerskunde,
afstudeerrichting mobiliteitsmanagement*

2011
2012

BEDRIJFSECONOMISCHE WETENSCHAPPEN

*master in de verkeerskunde: mobiliteitsmanagement
(Interfacultaire opleiding)*

Masterproef

*Modelling Public Transport Potential in Relation to Area
Characteristics*

Promotor :
Prof. dr. Gerhard WETS

Copromotor :
De heer Petrus VAN DER WAERDEN

Wannes Poesen

*Masterproef voorgedragen tot het bekomen van de graad van master in de verkeerskunde,
afstudeerrichting mobiliteitsmanagement*

Modelling Public Transport Potential in Relation to Area Characteristics

Wannes Poesen*
Prof. dr. Geert Wets
Peter van der Waerden

Transportation Research Institute (IMOB)
Hasselt University
Wetenschapspark 5, bus 6
B-3590 Diepenbeek
Belgium

Fax: +32 (0)11 269 111
Tel: +32 (0)11 269 199
IMOB@uhasselt.be

E-mail:
Wannes.Poesen@student.uhasselt.be
Geert.Wets@student.uhasselt.be
P.J.H.J.v.d.Waerden@tue.nl

*Corresponding author

Number of words: 10 782
Number of tables: 6
Number of figures: 0
Words counted: $10\,762 + 6 * 250 = 12\,282$ words

Paper submitted: June 1, 2012

ABSTRACT

The process of mode choice comprises different aspects. In this paper, models are developed linking area origin- and destination characteristics with reasons why people don't choose to take the bus to determine public transport (PT) potential. The need for this kind of research lies in the fact that PT providers need insights into PT potential to optimize their services. In order to know what characteristics to use, a brief overview of a literature is presented. This part focuses on area characteristics but also includes choice theories describing the process underlying mode choice behaviour. Hereafter, the ROVBECO research, for which the data used in this study was collected, is described. The independent variables for the models can be categorized into origin- (dwelling type and family status), destination- (employment size and building density) and travel mode (mode choice and speed) related characteristics. Two models determine the dependent variables: First, the mentioning of reasons of not taking the bus for a commuting is modelled. Second, if a reason was mentioned, the chance of the respondent switching from car/bike to bus for commuting trips when measures were to be taken to resolve the issue is modelled. The model type used in this study is binary logit. The main conclusion of the research is that the PT-potential model delivers interesting but rather limited results: building density seems to have the most significant influence followed by mode choice. Recommendations to increase model performance are the addition of more variables with more levels.

1 INTRODUCTION

In recent decades the market share of public transport has declined to a marginal proportion of the modal split. For instance, the share of PT in Flanders and The Netherlands is only around five per cent [1] [2]. The public interest of a well-functioning PT-system is evident, not only to mitigate the externalities of car traffic but also to raise the cost coverage of PT-providers. PT operators must invest in a better understanding of the mobility market and must be able to adapt to the varying needs of travellers in order to raise their cost coverage. To achieve these goals, better marketing tools need to be developed that don't tackle problems ad hoc, advertising campaigns and reductions on fees for instance, but do this in a structural manner [3].

The paper is structured as follows: in a first part the problem definition and research objective will be examined. Hereafter a literature review follows describing the mode choice process from two points of view: the utility maximization theory and the theory of planned behaviour. Factors influencing mode choice are described as well. In the fourth chapter the methodology is addressed together with the data preparation for the models. Hereafter the model results are discussed, followed by an application. In the final part of this paper a conclusion follows with recommendations for future research.

2 PROBLEM DEFINITION AND RESEARCH OBJECTIVE

From the foregoing it is clear that PT providers have not the required knowledge and tools to respond to the traffic and mobility market in a proper manner. The problem definition can be thus be described as follows: **Public transport providers need more knowledge and better tools in order to react to changes in the mobility market.**

In order to increase the knowledge of PT providers, one needs to identify those aspects that influence the reasoning process in choosing PT as a travel mode. The Regional Public Transport Accessibility Consumer-Oriented research (ROVBECO) which will be explained in further detail later was concerned with the issue of identifying PT potential as well. Both studies however posed some problems and limitations with the models being used. One of the problems was the limited number of three independent variables used which only covered the origin side of the respondent's trip. Another problem concerns the moderate performance of the model. A limitation of the model was that only car substitutability, meaning switching from car to PT, was concerned without taking into account possible reasons for this behaviour. The studies examined area or spatial characteristics to determine PT potential. The research question of this paper can be described as follows: **What characteristics of areas are important when examining public transport potential?**

This problem definition is important in order to reach the goal of a better understanding of modelling PT potential. At heart of this master thesis is the construction of the underlying models in order to increase the performance of the models and thus form a better basis for tools such as the GIS-based potential map. The main goals of this research can be described as: **To develop a model that links origin- and destination characteristics to reasons for not choosing public transport in order to determine public transport potential.**

A second goal of the study is to develop a complementary model that links again origin- and destination characteristics to switching behaviour to public transport.

With the indicated marginal proportion of PT-trips nowadays the potential for more PT-trips is substantial. This doesn't however mean that all non-users are willing to switch to PT. Some commuters might consider to take the bus but aren't able to do so with the current supply of PT services. Other commuters might not even take PT into consideration and are known as auto captives and are very difficult to convince in taking the bus. Studies on potential PT users show significant numbers of non-users (as high as fifty per cent) willing to switch to PT if adequate measures were to be taken, e.g. [4] [5] [6].

3 LITERATURE REVIEW

The above shows that PT providers lack adequate knowledge on the process of travel mode choice and mobility behaviour. One of the few attempts that were made to develop an adequate marketing tool is the so-called potential map. According to Van der Waerden et al. [7], this tool allows PT providers to locate the most likely customers of public transport by using area characteristics. The final output of the potential map tool is a geographic information system (GIS) based map that visually displays the locations of potential. The PT potential map is a way of geomarketing. In marketing terms, geomarketing is segmenting a population based on area characteristics. Central in geomarketing are the characteristics of a specific area [8]. Applied to transport and mobility, this concept thus implies the connection between travel behaviour and spatial patterns [9]. Geomarketing answers the question 'where' PT providers can optimize their services. With the use of geomarketing techniques like the potential map, PT-providers can adjust and optimise their services adequately in a given area for a given target population, e.g. adjusting the frequencies, route of the line, tariff measures and the location of stops.

Many aspects influence the complex process of mode choice. In this part an overview of literature is presented describing the mode choice process and factors influencing this process. Particular interest goes to the influence of area characteristics.

3.1 Mode Choice Process

When a researcher observes one's mode choice process he can observe but only a few factors influencing this process. One might think that mode choice is merely a simple process and a result of objective (e.g. area) characteristics leading to one's choice. But this deterministic view is too brusque; other unobservable subjective characteristics (perceptions) influence mode choice as well [10]. Taking into account the unobservable factors, the behavioural process might explain why different individuals make different choices in the same situation. As will be seen later, higher densities and a higher function mix will lead to more trips per PT but changes over individuals occur. Person A might choose not to use PT because he perceives the route between the bus stop and his work location as unsafe while person B perceives the same route as safe and thus chooses for PT. Behavioural theories contribute in explaining the differences over individuals in the process of mode choice. In literature one can find two main approaches to describe the mode choice process [11]. First, the utility maximization theory will be described. Second, a psychological behaviour theory, namely the Theory of Planned Behaviour, will be described.

3.1.1 Utility Maximization Theory

Discrete choice models are often used in modelling transport. These disaggregate models are based on observations of individuals. The theoretical background of these models is the utility maximization theory which states that people will always choose that option which maximises their net personal utility subject to legal, social, physical and budgetary constraints. This theory was developed by McFadden in the 1970's [12] [13]. The utility that a person attaches to an alternative mode is derived from the characteristics of that alternative. The higher the relative utility, the higher the chance one chooses the alternative.

A problem with this approach lies in the fact that the modeller doesn't possess complete information about all the elements considered by the individual making a choice, e.g. [12] [14]. The theory states that a person bases its choice on factors from which some are directly observable (e.g. area characteristics) and others are not (perceptions) [14]. The observable factors are labelled as x while the unobservable factors, the error term, are labelled ε . To conceive the mode choice one can put these variables into a behavioural process which is known in the function of $y = h(x, \varepsilon)$. The inference is that area characteristics per se will not directly lead to a certain choice. As stated before, the choice can thus be divided into an observable and unobservable part. The observed part is a function of measurable attributes while the unobservable part defines an individual's taste. The problem when using discrete choice models lies in the fact that the behaviour cannot be reproduced

exactly. There will always be an unobservable part that might be explained by theories from social psychology described next.

3.1.2 *Theory of Planned Behaviour*

One of many theories from social psychology dealing with choices is the theory of planned behaviour by Ajzen and Fishbein. The theory is an extension of the theory of reasoned action and states that underlying motivations and intentions lead to a specific behaviour [15]. The theory focuses on three kinds of beliefs: behaviour beliefs which influence attitudes, normative beliefs which influence subjective norms and control beliefs which influence perceived behavioural control [16].

One's attitude towards a specific behaviour is built up from behaviour beliefs. An attitude might be positive or negative where a negative attitude might lead to not perform the behaviour. The concept of attitude is equivalent to the utility maximization theory: one will show a more positive attitude towards an alternative that yields more utility. One assesses the attributes of an alternative but this theory differs from the utility maximization theory in that it takes also other factors into account like the subjective norm. This norm is built up from normative beliefs and indicates the extent to which an individual complies with referent individuals. These referent individuals have a certain opinion about the behaviour and will encourage or discourage this behaviour. The control beliefs influence perceived behavioural control and this in turn determines what factors facilitate or inhibit the behaviour [13].

These factors determine intention which in turn determines, together with perceived behavioural control, the behaviour. The more favourable the attitude, subjective norm and perceived behavioural control towards a behaviour, the more likely a behaviour will be performed. An example: when one's attitude towards the bus is positive as well as the subjective norm it might still be possible that this individual might not choose to take the bus because the bus simply does not operate on the desired hour. Thus the perceived behavioural control makes it impossible to perform the behaviour.

3.2 **Influencing Factors**

De Dios Ortúzar and Willumsen [12] make a distinction between a large set of factors influencing the process of mode choice: characteristics of the traveller (e.g. car ownership, possession of a driving license, household structure, income and residential density), characteristics of the trip (e.g. trip purpose, time of the day) and characteristics of the travel mode (e.g. travel time, costs, number of parking places, comfort and convenience, security, reliability).

Other factors influencing mode choice are area characteristics, e.g. [9] [15] [17]. A rather logical finding is that differences in mode choice can be found between city and rural residents [18]. Area characteristics influencing mode choice is a focus adopted only since a few years in marketing of PT [7] [19]. In this research areas are examined wherein potential public transit users can be found. It is not just the origin (the residential area in this context) of one's trip that might influence the mode choice but also the destination. A literature review on area characteristics influencing the use of PT concludes the following factors to have a significant influence on PT choice: the bigger the mix of functions (the ratio between housing, offices, shops and other facilities), the higher the number of PT trips [12] [20]. The higher the building density (number of buildings per hectare), the higher the higher the number of PT trips, e.g. [21] [22]. The presence of freestanding dwellings decreases the chance in taking PT while the presence of flats increases this chance [22] [23].

3.3 **Reflection**

When creating a theoretical framework for this research paper, it is important to notice that mode choice doesn't stem directly from objectively measurable characteristics. How an individual perceives a location or a travel mode influences the choice as well. The utility maximization theory takes only observable factors into account and collects unobservable behavioural factors into an unobservable term. Together with the theory of planned behaviour, these theories assume that the individual makes a well-considered choice. But the theory of planned behaviour looks more at subjective factors. Handy [13] states that theories from transport geography, such as the utility maximization theory,

determines the mechanism of behaviour while theories from social psychology, such as the theory of planned behaviour, determines those factors that have an influence on behaviour. The former theory describes the relationship between factors and mode choice while the latter theory goes deeper into the reasons of the choice made, i.e. the ‘why’-question.

Both theories described assume that individuals make a well-considered choice. However, when an individual performs the same behaviour over and over again, the choice will be influenced particularly by a habit instead of other factors, e.g. [15] [24] [25]. A habit is instigated by a specific goal-directed state of mind in the presence of triggering stimulus cues [24]. The initial choice for a specific travel mode for a specific goal will result from a reasoned process by taking different sorts of factors into account but after some repetitions, the process will disappear and only the habit will persist. A distinction can be made between a strong and a weak habit. Someone with a strong habit doesn’t take much information into consideration in the mode choice process while someone with a weak habit considers more information prior to the choice. Since the activity pattern of employed people is relatively stable, the above leads to the conclusion that not many factors are considered before choosing a particular travel mode [26] [27].

It is clear that not a single theory is capable of clarifying mode choice behaviour fully. Arentze et al. [28] stated that more complex theories and methods are needed to conceive mode choice. A possible solution is to expand or combine existing theories [13]. An example of casting this study into a broader theoretical framework could be to expand the perceived behavioural control in the theory of planned behaviour by inquiring area characteristics. These results could then shed light on what factors facilitate or inhibit mode choice. Although a broader perspective is necessary in this research, this framework will be limited by the available data sources as will be seen later. In this paper it is assumed that respondents follow some kind of mode choice process rather than relying on a habit. The mode choice process is based on the utility maximization theory.

4 RESEARCH DESIGN

This research builds on the proceedings of the before mentioned studies. The focus here lays on public transportation, in particular the bus. Commuting trips is the only trip motive under consideration, even though commuting trips account only for a small fraction of all trips [1]. The focus on commuting trips is straightforward since these trips are more concentrated in time and space. Of particular interest is the ROVBECO research. This part is concerned with a further exploration of the dataset and the according ROVBECO research for which the data was initially gathered. Data preparation and model choice is described in further detail before continuing with the analysis.

4.1 Exploration of the Dataset

In the modelling phase, the dataset of the ROVBECO research has been used. Because of the importance of this dataset, further information on the research is appropriate. The ROVBECO research [17] [29] was conducted in the SRE-zone in The Netherlands. The SRE is a regional cooperation between 21 municipalities in the southeast of the province of Northern Brabant and deals with topics exceeding the municipal boundaries, e.g. regional transportation, spatial policy and socio-economic affairs [30]. SRE is responsible for the organization of PT in the SRE-zone. Main purpose of the ROVBECO research was to create a tool with which PT potential can be identified and to identify underlying reasons for not using PT (bus only).

The ROVBECO research can be divided into three main parts. The first part concerned the potential of PT trips. The potential to use the bus was defined (and limited to the car only) as the chance of a car user to switch to the bus. The second part dealt with measures that could be taken to convince car users to switch to PT. Of major concern were the questions in what manner a non-user would react to an improvement of a certain issue (a measure) that led him initially to not use the bus: switch or don’t switch to the bus. The third part concerned the accessibility of those areas that were labelled, through the use of the model, as areas with a high potential for bus trips. Respondents could check one or more reasons for not using the bus from a list of sixteen possibilities. This list was

predefined meaning that respondents couldn't state their own reasons. An overview of the reasons is presented in table 1.

TABLE 1 Overview of the Reasons in Not Taking the Bus

The bus stop is too far from my home (ma)	Bus services are not provided early enough (mi)
There is no bus stop near my destination (mb)	Bus services are not provided late enough (mj)
The bus frequency is too low (mc)	The connection between bus and train is poor (mk)
There is no direct connection to the desired destination (md)	The buses are too crowded (ml)
Travel time by bus is longer than by bike (mef)	Information about schedules is insufficient (mm)
Travel time by bus is longer than by car (mef)	The number of seats is too low (mn)
It's difficult to carry stuff along (mg)	The bus (driver included) is not customer friendly (mo)
Travelling by bus is too expensive (mh)	It is not safe enough on the bus (mp)

The dataset consists of an internet-based survey, geographic data obtained from a data warehouse Bridgis (www.bridgis.nl) and data on the transportation networks. For the internet-based survey, around 33,500 invites were distributed among inhabitants of the SRE. Around 1,400 questionnaires were completed. The questionnaire contained questions about travel behaviour, socio-economic characteristics, the perception on bus facilities and the assessment of some relevant characteristics of the bus system in the region. Only three major travel modes were distinguished in the questionnaire. The revealed-preference cross-sectional data from the survey did not only contain commuting trips but also recreational and shopping trips. These trips were eliminated yielding a modal split in the SRE-zone for commuting trips of around 41% car, 31% bicycle and 28% bus. Geographic data from Bridgis contains the number of households, the number of dwellings, the number of persons, the predominant type of dwelling, the family status and the employment size of companies per 6-position zip code area. A 6-position zip code area covers about street. This data was added to the survey together with information of the transportation networks. Since this research focuses on potential bus trips, respondents who already travelled by bus were eliminated.

The final car replacement model in ROVBECO is a binary logistic regression model with three independent variables, significant at a confidence level of 95%. All variables contain two attribute levels and are effect coded. The first variable, type of dwelling, contains the levels half-detached and detached homes versus terraced houses and apartments. It was found that the presence of detached homes lowers the chance in switching to use the bus while the presence of flats increased the chance. The second variable, bus frequency, contains the levels no frequency and one or more buses per hour. A logical effect was found in that an increase in bus frequency increases the chance of switching to the bus. The third and final variable, respondents' origin, contains the levels residing in Eindhoven and residing outside Eindhoven. The result here was surprising: respondents outside Eindhoven were more prone in switching to PT than respondents residing in Eindhoven. This could be down to the fact that respondents in Eindhoven use the bus already more than those outside Eindhoven.

The limitation of the ROVBECO research lies in the fact that trips are only considered from the origin side: the dwelling area. Since a trip is a movement between an origin and a destination, one could assume the destination having a significant influence as well on mode choice, e.g. [31] [32] [33]. Indeed, if no bus stop is present or near the destination, the chance of taking the bus will deteriorate, e.g. [34]. Since the available data also contains information on the respondents' destinations and the relationship between origin and destinations wasn't directly analysed in ROVBECO, this dataset is of particular interest and deems ideal to this research (cfr. research goal).

The final dataset for this research further eliminates uncompleted queries and rounds up to 432 useful queries. The modal split of this dataset is 54.6% car and 45.4% bicycle. Compared to ROVBECO where bicycle users were excluded from the model as potential for PT, in this research bicycle users are included. This increases the number of respondents. After comparing the dataset to average Dutch figures from the Central Bureau of Statistics, It can be concluded that the dataset is broadly in line with the average Dutch figures. Some general figures about the 432 respondents are: the average age is 49 years old, compared to the Dutch average of 40.3 this number is acceptable.

Youngsters are slightly underrepresented while elderly are overrepresented. 60% of the respondents are male while 40% are female; this ratio is acceptable compared to the 49.5-50.5 ratio of The Netherlands. The level of education is slightly skewed towards a higher percentage than average of highly educated respondents: 45% compared to the Dutch average of 34%. Middle schooled respondents sum to 50% compared to 61% on average and primary education sums to 5% corresponding to the average of 5%. The number of young respondents in possession of a driving license is rather low compared to the Dutch average. Other age categories are in accordance to the CBS figures [35]. Over 91% of the respondents possess a driving license; this is slightly higher than the 81.3% average. The number of cars per household is on average 1.31 while about one in four has a bus subscription. The nationwide figures point out that around 60% of all trips is shorter than 6km while in the dataset this number is a mere 19%. The rest of the trips are spread out rather equally over the other categories. A difference between the nationwide and ROVBECO trip duration can be noticed as well. Although the different categories of variables being considered are sometimes slightly skewed compared to the CBS figures, the dataset can be considered representative.

4.2 Data Preparation

Before estimating the models it is important to define the variables and the manner in which they are specified. Central to this research is PT potential and the reasoning for (not) switching to bus. It is important to highlight this difference with the previous researches. Unlike the ROVBECO research, reasons for not taking the bus will be taken into consideration. These are the sixteen reasons from table 1 with which respondents indicated why they didn't take the bus. First the dependent variables will be described.

The process of data preparation was done by using IBM SPSS Statistics and Microsoft Excel. First the questionnaire data, together with the Bridgis data, needed to be effect coded and converted to a format suitable for NLOGIT 3.0, the program in which the models will be estimated [36].

4.2.1 Dependent Variables

The analysis consists of two models and will be based on two different dependent variables. In the first model of the analysis, the chance of a reason to be mentioned by a respondent will be examined. In the second model the potential switch to PT is examined.

These two model lead to different dependent variables. In the first case, the dependent variable is described as the chance of a reason to be mentioned by a respondent; that is one of the sixteen reasons from table 1. In the second case the dependent variable is described as the chance of a respondent to switch to the bus if the issue were to be resolved. Both models make use of the same independent variables. The independent variables used for the analysis can be categorized into origin characteristics (dwelling type and family status), destination characteristics (employment size and building density), a travel mode variable (speed), mode choice (car or bicycle user) and the reasons for not taking the bus themselves. All independent variables consist of two attribute levels. These variables will now be described in more detail.

4.2.2 Independent Variables

Another difference with the ROVBECO research is the addition of characteristics of the destination side, next to the already present characteristics of the origin side. Another addition is the use of a transport mode characteristic. The data for the variables comes from the ROVBECO dataset and the corresponding dataset by Bridgis. Data from Bridgis is provided at a detail level of 6-position zip codes. However, since respondents only indicated their destination at a 4-position zip code level, the Bridgis data is aggregated from a 6-position level to a 4-position zip code level. A 4-position zip code level covers about a neighbourhood. This aggregation means a loss in the level of detail of the data. However, the loss seems rather small.

Mode Choice Since bicyclists already use a very sustainable mode of transport and car users don't, it is not unthinkable of to assume a modal shift to occur more quickly amongst bicyclists. A meta-analysis by Van den Bergh et al. [37] showed that an experiment to raise the number of PT-users did indeed success, however the number of car users switching to PT was relatively low compared to the number of bicyclists switching to PT. This indicates that bicyclists are more prone to switch to PT [38]. This variable contains two levels or possibilities: the respondent uses a car or a bicycle for its trip in the current situation. The mentioning of reasons, as well as the switch to PT, might thus depend on the respondents' current mode choice.

Speed Instead of using the travel time ratio like the ROVBECO research, speed is used instead. There is a direct link between travel time ratio and mode choice: the higher the ratio, the lower the chance that a respondent will choose the bus and vice versa. However, due to the lack of individual travel time ratios per respondent, speed is used. According to Corpuz [39] speed affects the mode choice as well. Slow speeds may indicate congestion. This encourages travellers to switch to PT. The effect is however rather low [39]. Two levels are noticeable in this variable: a slow speed and a high speed. It might thus be possible that respondents with a low speed mention other reasons than respondents with a high speed. This assumption holds for the switch to bus as well. To determine whether the speed of a respondent is low or high, reference data is needed. Because of data availability the Flanders travel behaviour research was used (OVG) [1]. This data is broadly in line with the Dutch figures. This data contains average speeds for every travel mode (e.g. car, bus and bicycle) and motive (e.g. commuting, recreational and shopping) per distance category. For instance: If a respondent indicates to have a commuting distance of 8 kilometres and does this in 10 minutes by car, then, according to the average speed indicated by the OVG, the speed is high. If the same respondent would have indicated that the trip took not 10 but 12 minutes, then the speed is low [1].

Dwelling Type The ROVBECO research indicated that the type of dwelling has a significant influence on mode choice, as described above [17]. Respondents residing in flats could mention certain reasons while respondents residing in detached houses won't. For this reason the variable is incorporated here. The Bridgis data contained detailed data on the number of dwelling types. Fourteen types were considered. In the SRE the most common dwelling types are detached houses and terraced houses. Other types of dwellings do occur but are negligible. The dominant dwelling type is the type of dwelling that occurs most frequent in a certain zip code area. Two possible dwelling types are incorporated in this variable: detached houses on the one hand, duplex and terraced houses on the other.

Family Status A number of studies showed that travel behaviour and mode choice in households is influenced by the household structure, e.g. [12] [40] [41]. The dominant family status is incorporated as was in the ROVBECO research, i.e. the most common type of family status in a zip code area. The variable has got two levels: families without children and families with children. The main idea is that families with children show a different travel behaviour than those without children.

Employment Size Employment size will be used as a proxy for the type of working destination. If the average employment size of an area is rather high, then it is assumed that the area is a central business district (CBD). According to de Abreu e Silva et al. [31] and Frank & Pivo [32] dense employment will reduce the number of trips made by car and increase the number of PT trips. This means that those respondents working in a CBD for instance might be more prone to switch to the bus than those working in other locations. Again data of Bridgis was used to determine the employment size of the different areas. By dividing the destinations into high or low employment sized areas (the two levels of this variable), 7 high employment areas are found. 6 out of 7 are situated in Eindhoven and 1 is situated in Helmond. This was done by looking at the ratio of the number of employees per squared kilometre.

Building Density Multiple studies show the positive effect of building density on PT ridership, e.g. [21] [42]. The higher the density, the higher the chance of PT-trips and the lower car use. This variable has got two levels: high and low building density. Instead of using the number of buildings, the number of dwellings is used since the Bridgis dataset only contains this number. The threshold for being a high density area is set at 160 dwellings per square kilometre. This is slightly lower than the Dutch average of a 190 dwellings per square kilometres [43]. Areas with more than 160 dwellings are defined as areas with a high building density while areas with less than 160 dwellings are defined as areas with a low building density.

4.2.3 Specification of the Variables

The variables are effect coded and are reason specific. This means that the utility of the reasons can be determined specific for one of the sixteen reasons. Effects coding has the advantage over dummy coding in that the effect of a variable isn't measured against a reference category [36]. This means that the effect can be derived directly. The specification was done as follows:

TABLE 2 Specification of the Independent Variables

	Variable	Code Name	Levels	Coding
Travel Mode	Mode Choice	v	Car	-1
			Bicycle	+1
	Speed	s	Fast	-1
			Slow	+1
Origin	Dwelling Type	dw	Duplex & Terraced	-1
			Detached	+1
	Family Status	dg	Without Children	-1
			With Children	+1
Destination	Employment Size	b	Low	-1
			High	+1
	Building Density	w	Low	-1
			High	+1

The naming of the reasons in the model is indicated between brackets in table 1.

4.2.4 Model Selection

In the literature review it was stated that social psychology theories give more insight into the specific reasoning in the process of mode choice. However, the questionnaire of the ROVBECO research wasn't conceived as pure attitude survey. This leads to the choice of the utility maximization theory as a theoretical framework for the model choice. The theoretical framework, the type of data and the specification lead to the conclusion that a discrete choice model will be used [36]. As was mentioned before; the data is revealed-preference cross-sectional.

The model type to analyse the relation between mode choice and characteristics of origin and destination will be the multinomial logit model. This model specification predicts, by means of defined parameters, the chance that a certain alternative will be chosen. This chance is determined by computing the utility of each alternative through an expected value and an error term. To derive the utility of an alternative, one uses the following formula:

$$U_i = V_i + \varepsilon_i$$

With U_i being the utility of alternative i ; V_i being the expected value of alternative i based on the independent variables and ε_i the error term of alternative i . The error term is necessary since different individuals make different choices even in the same situation. The error term thus describes how much the individual deviates from the mean V_i . The error term contains, as stated before, unobservable factors. The error term has a probability distribution with an expected value of zero [14] [36]. Since other alternatives are involved, the chance of choosing one alternative needs to be calculated by taking into count the other alternatives. This is done by the following formula:

$$P(i) = \frac{\exp(U_i)}{\sum \exp(U)}$$

With $P(i)$ the chance of choosing alternative i , U_i being the utility of alternative i and $\sum(U)$ is the sum of all utilities. As described before, two dependent variables will be used. This means that two models will be estimated. The first model describes the chance of a reason in not taking the bus to be mentioned. These reasons are the sixteen possible reasons described above in table 1. The dependent variable here is the chance of a reason to occur. The second model describes the chance to switch to the bus when the issue is resolved by taking appropriate measures. This model thus considers the willingness to switch.

The two models consider only two alternatives. The first type considers the chance of the specific reason to occur; the alternatives are the reason being mentioned and the reason not being mentioned. And the second type considers the chance whether a respondent will switch to the bus or not; the alternatives here are to switch or not to switch. All these chances are examined per reason. Because only two alternatives are considered in the models, the former formula will become a binary logit model, described by the following formula:

$$P(i) = \frac{\exp(U_i)}{\exp(U_i) + 1}$$

In the following part, the model results will be discussed.

5 RESULTS

The modelling process started with several trial models. These models contained only a few variables in order to understand the data. More variables were added systematically and different model specifications were tested in order to arrive at the final models. In deriving the final models in the modelling process, a few requirements were taken into account. One is that the variables do not suffer from multicollinearity, meaning that the independent variables are not strongly correlated with each other. This requirement holds true for all variables. Second, LIMDEP poses a limitation to the number of variables being used. The number of variables in the final models however didn't cause any problem. Finally, a sufficiently large number of observations is needed per variable. With a dataset of 432 respondents, this number seems sufficient for the final models. Next, the results of the models will be discussed.

A first look at the data shows us that 22% of the respondents are auto captives since these people indicate to not switch to the bus even if all the issues they indicated were to be resolved. This figure doesn't correspond to the 47% found in the American research [4] but does correspond to another study carried out in The Netherlands [5]. The difference might be explained by cultural differences, e.g. [6]. However, the figure shows that 78% of the respondents might be susceptible to use the bus if those issues were to be resolved. This number indicates that the data forms a solid basis to analyse and derive adequate results.

A discussion of both models follows. First, the chance of a reason being mentioned will be analysed. Secondly, the chance of the respondent switching to PT if the mentioned reason is resolved by taking adequate measures will be analysed. During the modelling process, insignificant variables were removed resulting in a model in which all variables are significant at a confidence level of 95%. As stated before, the utility that a respondent assigns to a specific alternative is defined as $U_i = V_i + \varepsilon_i$. Since only V_i is represented by observable factors (the independent variables), we can only make a statement about the form of this factor. This factor takes on the form of:

$$V_i = \beta_{0i} + \beta_{1i}(fX_{1i}) + \dots + \beta_{Ki}f(X_{Ki})$$

Where β_{1i} is the parameter associated with attribute X_1 and alternative i . And β_{0i} is the alternative-specific constant (ASC). This is a parameter that is not associated with any of the observed and measured attributes and represents the unobserved sources of utility [36].

5.1 Estimation of Model I

This model determines the chance of a reason, as listed in table 1, to be mentioned. The output of the model is presented below. The independent variables to this model are: mode choice, building density, speed, dwelling type and employment size. During the modelling process, no significant relationship with family status was found between the chance of a reason to be mentioned or not.

TABLE 3 Output of Model I

Groups	Variable	Coefficient	Standard Error	b/St. Er.	P[Z >z]
Alternative Specific Constants (ASC)	MA	-1.786368621	.13722877	-13.017	.0000
	MB	-2.142286101	.18381822	-11.654	.0000
	MC	-0.610909082	.10074907	-6.064	.0000
	MEF	1.609437912	.22074255	7.291	.0000
	MG	-2.745421547	.26736001	-10.269	.0000
	MH	-0.975934518	.12064543	-8.089	.0000
	MI	-1.526708848	.13382276	-11.408	.0000
	MJ	-1.529467108	.13315685	-11.486	.0000
	MK	-2.460091481	.17867380	-13.769	.0000
	ML	-1.42310027	.12388993	-11.487	.0000
	MM	-2.103095733	.15451213	-13.611	.0000
	MN	-1.967244292	.14665077	-13.414	.0000
	MO	-2.631640732	.19226063	-13.688	.0000
MP	-3.782107403	.40238125	-9.399	.0000	
Mode Choice	VD	-0.330629414	.10184580	-3.246	.0012
	VG	-1.125779464	.26736001	-4.211	.0000
	VI	-0.29613745	.14258464	-2.077	.0378
	VP	-0.89911645	.32206315	-2.792	.0052
Building Density	WEF	-0.78845736	.22074255	-3.572	.0004
Speed	SI	-0.317610132	.13823616	-2.298	.0216
Dwelling Type	DWB	-0.363625705	.17900564	-2.031	.0422
	DWH	-0.257597089	.12064543	-2.135	.0327
	DWJ	0.338481499	.13315685	2.542	.0110
	DWP	-0.649207352	.32343554	-2.007	.0447
Employment Size	BB	-0.341488062	.15075885	-2.265	.0235
	BD	-0.51295097	.10184580	-5.037	.0000
	BI	-0.274504755	.13303470	-2.063	.0391
	BL	0.336910501	.12388993	2.719	.0065

To understand the table better, first some explanation on how to read the table above is given. The first column depicts the name of the independent variables and the ASC's. The second column shows the specific reason for which the variable holds. The third column shows the parameter associated with the variable. To determine whether a parameter is statistically significant, the standard error (fourth column) is needed to perform the Wald-statistic. The result of this statistic is shown in the fifth column. At a confidence level of 95%, the critical value of the Wald-statistic lies at 1.96. For a parameter to be statistically significant, the absolute value of the Wald-statistics needs to be higher than 1.96. This requirement applies to every parameter in the table above meaning that all variables

are statistically significant. The model has got a R^2 value of 0.39257 implying an indicative fit for binary logit models [44]. The value of R^2 ranges between 0 and 1 with 1 being the optimum.

It was described above that the ASC is a parameter that represents the unobserved sources of utility. Looking at table 3 it can be seen that reason *a*, is only explained by unobservable attributes. Therefore it is not possible to conclude whether some area characteristics have an influence on the chance of the reasons being mentioned. But what can be concluded is that regardless of the situation, these reasons are important. The reasons that are only explained by the ASC's are: the bus frequency is too low (reason *c*); the connection between bus and train is poor (reason *k*); information about schedules is insufficient (reason *m*); the number of seats is too low (reason *n*); the bus (driver included) is not customer friendly (reason *o*). Other reasons are explained by the ASC and/or other variables, describing observable utility. Table 3 summarizes all the significant variables of the model.

Comparing the different ASC's (*ma* through *mp*), it can be shown that reasons *e* and *f* (travel time by bus is longer than by car/bicycle) are the most frequently mentioned reasons by all respondents in the questionnaire. Reason *p* was the least mentioned reason (it is not safe enough on the bus), hence the lowest value.

This leaves us to nine other reasons that can be explained by the independent variables. To determine the actual chance of a reason to be mentioned, one uses the earlier described formulas. For instance, the chance of reason *e* and *f* to be mentioned, we first calculate the utility associated with this reason. $U_i = V_i + \varepsilon_i$ where $V_i = \beta_{0i} + \beta_{1i}(fX_{1i}) + \dots + \beta_{Ki}f(X_{Ki})$. V_{ef} becomes $1.6094 - 0.7885 * BuildingDensity$.

Since all variables are effects coded at two levels, as is building density, by filling in -1 or +1 we obtain a different effect caused by the variable building density.

V_{ef} becomes 0.8209 in the case of a high density (coded as +1) and 2.3979 in the case of a low density (coded as -1).

To compute the chance of the reason to be mentioned when building density is high, one simply fills in the formula:

$$P(ef) = \frac{\exp(0.8209)}{\exp(0.8209) + 1} = 0.6944$$

The chance of this reason not to be mentioned in dense areas is $1 - 0.6944 = 0.3056$. The chance of the reason to be mentioned becomes 0.9167 when building density is low. Further calculations were made by using Excel. Except for reasons *d* and *ef*, chances of a reason to be mentioned were rather low (between 2 and 35%). Yet some conclusions can be taken about the chance a reason was mentioned:

- Reason *b*, There is no bus stop near my destination, is influenced mostly by respondents working in a low employment zone. This finding is stems with the literature;
- Reason *d*, There is no direct connection to the desired destination, is explained completely by mode choice and building density since the alternative-specific constant was not significant. The chance of this reason to be mentioned was highest among car users working in a low employment zone. The connection with low employment seems clear; however the relation with mode choice is not clear.
- Reason *ef*, Travel time by bus is longer than by bike or car, is mostly influenced by low building density destinations. This relation may be explained by slow bus lines with a high number of stops residing in low building density areas.
- Reason *g*, It's difficult to carry stuff along, is mostly influenced by car users. This seems logic since bicyclists are used to carry along things on their bikes which isn't always too easy.
- Reason *h*, Travelling by bus is too expensive, occurred more among respondents living in terraced houses than detached houses. A relationship might be found in socio-economic characteristics. Perhaps respondents residing in detached houses have a larger budget and hence find the bus fares rather inexpensive.

- Reason *i*, Bus services are not provided early enough, occurred the most among respondents using the car, having a high commuting speed and working in a low employment area. No clear connection can be found.
- Reason *j*, Bus services are not provided late enough, is mostly influenced by respondents residing in detached houses. An explanation could be the fact that detached houses are situated in non-urban areas where bus services stop earlier compared to urban areas.
- Reason *l*, The buses are too crowded, was to be more of a problem amongst respondents working in high employment zones. Since more PT trips are made within these zones (see literature above) and thus buses are more occupied, this result seems logical.
- Reason *p*, It is not safe enough on the bus, is mostly influenced by respondents living in terraced houses and driving a car. No clear relation can be found.

Most of the above conclusions seem logical and stem with the literature. However no direct explanations for the relationships of reasons *d*, *i* and *p* on the one hand and the variables on the other are found. It may be that reason *i* is influenced most by car users that need to travel long distances to work. In order to be on time and to avoid congestion, this might be an explanation. For reason *p* it seems that the opposite seems more logical: one could assume those residing in terraced houses live in more urban areas where social control occurs more than in rural areas. Because of more social control, one could assume that respondents residing in urban areas to find it safer to use the bus than those living in rural areas. Apparently, the opposite holds.

5.2 Model I Application

With the use of the first model one determines the chance of a reason being mentioned by a respondent or not. As was seen above, this chance might be influenced by some variables i.e. the mentioning of some reasons are area specific. An example is created to clarify the findings.

Consider an area with a population of 100 inhabitants who don't use the bus in the current situation. The dominant type of dwelling in this area is an apartment building. Other characteristics need to be considered person per person. Since family status wasn't a significant variable, this variable isn't considered here, i.e. it makes no difference in the choice of travel mode if one has got any children. The 100 inhabitants all have different characteristics. However, some of the above variables hold for multiple inhabitants. A distribution can be seen: 50 inhabitants work in a central business district (CBD), implicating a high building density and a high employment size. Of these 50 inhabitants, 15 commute by car and have a relatively slow speed, 15 commute by car and have a relatively high speed, 10 commute by bicycle and have a relatively slow speed while the other 10 commute by bicycle as well but have a relatively high speed. Of the remaining 50, 25 work at an industrial site, implicating a low building density and a high employment size (e.g. offices) while the other 25 work at another industrial site, implicating again a low building density but this time employment size is low (e.g. warehouse and storage facilities). Of the 25 working at an industrial site with high a employment size, 5 commute by car and travel at relatively high speeds, 10 commute by car at relatively low speeds, 5 commute by bicycle at relatively low speeds and 5 commute by bicycle at relatively high speeds. Of the remaining 25, the following distribution can be found: 5 commute by car at relatively low speeds, 5 commute by car at relatively high speeds, 5 commute by bicycle at relatively low speeds and 10 commute by bicycle at relatively high speeds.

The chances of respondents mentioning the reasons in not taking PT, were calculated by taking into account the different characteristics. First, a new dataset was created with a fictional list of 100 persons. All persons were then given the different effect codes according to the variables described above. With the parameters from the model, the utilities were determined. After this, all the chances were calculated per person. When aggregating the different chances for the entire 100 inhabitants, we find the following chances for every reason in table 4.

TABLE 4 Chances for Mentioning a Reason for a Fictive Example

	Reason															
	a	b	c	d	ef	g	h	i	j	k	l	m	n	o	p	
Chance for Mentioning in %	14	13	35	45	81	10	33	17	13	8	20	11	12	7	6	

In table 4 the different chances are based on the parameters from model I. measures *a*, *c*, *k*, *m*, *n* and *o* hold for all are only explained by the ASC's. As can be seen in table 3, other reasons are explained by more variables meaning that the chance of mentioning a reason becomes area specific. For the area described above, it can be seen that the chance of mentioning measure *e* and *f* are highest, followed by reason *d*. Since 100 inhabitants were considered, the percentages in table 4 stem with the number of inhabitants that will mention a reason: 81 people will mention reasons *e* and *f*, 45 will mention reason *d* and so on.

A remark with this example is that information on the different variables is needed. Data warehouses, like Bridgis, can provide the necessary data to specify the variables dwelling type, employment size and building density. The variables mode choice and speed can be derived from a national travel behaviour research like the OVG. In the example above, assumptions were made on these variables.

5.3 Estimation of Model II

The second model will determine when a respondent is willing to switch to PT. As was the case for model I, the same independent variables hold for model II. The variables are: mode choice, building density, speed, dwelling type and employment size. The model output is summarized in table 5:

TABLE 5 Output of Model II

Groups	Variable	Coefficient	Standard Error	b/St. Er.	P[Z >z]
Alternative Specific Constants (ASC)	MD	0.7932	0.1495	5.305	.0000
	MEF	1.0153	0.1350	7.519	.0000
	MH	1.7674	0.2625	6.733	.0000
	MI	1.0159	0.2547	3.989	.0001
	MJ	0.8008	0.2566	3.120	.0018
	ML	0.5875	0.2355	2.494	.0126
Mode Choice	MN	1.1239	0.3193	3.520	.0004
	VD	-0.4773	0.1495	-3.192	.0014
	VEF	-0.6861	0.1350	-5.081	.0000
Building Density	VH	-0.5352	0.2625	-2.039	.0414
	WC	0.3455	0.1646	2.098	.0359
	WK	1.3499	0.4241	3.183	.0015
Speed	WM	0.7577	0.3129	2.421	.0155
	SO	1.1451	0.4339	2.639	.0083
Dwelling Type	DWB	1.2763	0.3265	3.909	.0001
Employment Size	BL	0.4623	0.2355	1.963	.0497

The same explanation on how to read table 3 holds for table 4: The first column depicts the name of the independent variables and the ASC's. The second column shows the specific reason for which the variable holds. The third column shows the parameter associated with the variable. The result of the Wald-statistic is shown in the fifth column and the values of the fourth column were used to calculate this value. Again, all variables are significant at a confidence level of 95%. The model has got a R^2 value of 0.16653, a moderate fit (Louviere et al., 2000). Note that this value cannot

be compared to the previous model since the models have different dependent variables. This model now looks at the chance of a respondent switching to PT if the issue is resolved by taking adequate measures. As for the first model, the different chances for the PT switch are calculated.

The following reasons are not related to any independent variables and are solely explained by the ASC's: reasons *i* and *j*. However, if these were to be resolved, one could expect the chance of a respondent, who mentioned the corresponding reason, willing to switch to the bus to be the same in any condition. If bus services were provided early and late enough anywhere, there would be a potential of respondents willing to switch to PT. Reasons *a*, *g*, *n* and *p* were not significant. Other reasons were explained by the ASC and/or an independent variable and are thus area/context specific. An overview:

- Reasons related to mode choice are d, ef and h:
 - If a direct connection to a destination would exist, more car users would switch than bicyclists.
 - If the travel time of the bus were to be reduced compared to the travel time of the car and bicycle, a significant amount of respondents would switch to the bus
 - If the bus was less expensive, more car users would switch than bicyclists. However, the amounts are again significant.
- Reasons related to building density are c, k and m:
 - If the bus frequency was higher, more respondents working in high density areas would switch to the bus compared to those working in low density areas.
 - A significant difference can be noticed between those respondents working in a low density area compared to those working in a high density area when it comes to the connection between bus and train. When this connection was better, significantly more respondents working in high density areas would switch.
 - If the number of seats was higher, again more respondents working in high density areas would switch compared to those working in low density areas.
- Reason related to dwelling type is b: if the bus stop would be closer to the work location, then significantly more respondents living in a detached house would switch than those living in a terraced home or apartment.
- Reason related to employment size is l: more respondents working in high employment zones would switch to the bus these would be less crowded than those working in low employment zones.
- Reason related to speed is o: if the bus and its bus driver would be more customer friendly, then more people now travelling at higher speeds would switch to bus than those travelling at slower speeds.

Some reasons are only explained by unobserved factors, i.e. ASC's. This means that no specific area characteristics are influencing these reasons. This implicates that PT providers can tackle these issues rather easily in that they do not need to take into account specific environments.

5.4 Model II Application

As was shown with the first model, the second model can be applied to the same example as well. Consider the same area of 100 inhabitants with the same area characteristics. The distribution of the work places and other variables remains the same. Again, by using the dataset from the first model example and applying the parameters from the second model, one can calculate the number of inhabitants willing to switch to PT.

TABLE 6 Number of Inhabitants Willing to Switch

	Reasons										
	b	c	d	ef	h	i	j	k	l	m	o
Switching to PT in %	22	50	69	73	85	73	69	50	69	73	50

From table 6 it can be seen that, depending on the before described characteristics, that the highest potential lies in the fact that if measures were to be taken to resolve reason *h*, 85 people are willing to switch to the bus. If measures were to be taken to resolve other reasons as *ef*, *m* and *I*, 73 people will switch. Depending on the reason, one sees a potential of 22% to 85% for this particular case, i.e. depending on the variables. This potential stems with the general literature, e.g. [4] [5] [6]. Again, not every switch to PT is area-specific; if measures were to be taken in order to resolve reasons *i* and *j*, then in every situation 73 respectively 69 people would switch to PT.

6 CONCLUSION

In this paper two models were developed in order to define the potential of public transport. The dependent variable for the first model was the chance of a respondent mentioning a reason for not taking the bus. The dependent variable for the second model was the chance of a respondent switching to the bus when the issue (the reason mentioned in the first model) were to be resolved by adequate measures. The independent variables for both models were origin characteristics, destination characteristics, a travel mode variable and the current mode choice. It can be seen that even if a reason was mentioned in the first model, it didn't necessarily imply that the respondent would switch to PT in the second model. Of course, only those willing to switch to PT in the second model mentioned the reasons in the first model.

In the first model, only one reason had in advance a significant chance of being mentioned. This can be derived from the high ASC value associated with this reason. Other reasons were less likely to be mentioned by the respondents. However, it was found that mostly mode choice, dwelling type and employment size influenced the mentioning of the reasons. The model itself performed rather well. An example showed the application of the model on a fictive example.

In the second model the chance of a respondent switching to bus when a reason was mentioned was calculated. This model performed rather medial. In the same fictive example, the model was applied to demonstrate its operation. The potential of some reasons can be applied to any area without taking into account different characteristics. These reasons are explained by the ASC's only. Other reasons, however, are nuanced per area. Building density and mode choice seemed to have the most influence on the PT potential.

7 DISCUSSION AND FUTURE RESEARCH

The goals of this research were: To develop a model that links origin- and destination characteristics to reasons for not choosing public transport in order to determine public transport potential. And to develop a model that uses the same independent variables but now the dependent variable was to determine switching behaviour to public transport of those respondents who mentioned a reason for not choosing PT. This research is an addition to previous attempts in modelling PT potential in that destination characteristics and a travel mode characteristic were taken into account as well. A further addition to previous research was that specific reasoning for not taking PT and PT potential was taken into consideration as well. This is a handy feature for PT providers in that they can apply concrete measures in those areas where specific reasons hold potential for PT use. Although the models screen areas for reasons in not taking PT and PT potential in a global manner, with the use of the present models PT providers can screen different areas with different characteristics for potential. This was presented in the model applications. Since reasons for not taking the bus were incorporated and PT potential was based on these reasons, PT providers are able to take adequate measures more easily.

Since this research considered the relation between origin- and destination characteristics only in a rudimentary way, this is where the limitations of this research lie. The models delivered rather limited results in that potential was nuanced at best by two variables. All independent variables contained only two levels. When characterizing different areas with the independent variables, this implies that many areas will be characterized in the same manner even though variation between similar areas exists. When applying these models to other areas, one needs to keep in mind the limitations of this research. Another limitation might be the specification of the dependent variable in the second model. Now the outcome of the variable is: the respondent switches to PT or not. This

definition might seem too stringent or definite, i.e. there's no room for those who doubt to switch. The model showed for some reasons indeed a high number of potential, but it might well be that a significant share of those respondents indicating to switch won't switch if an issue is resolved because other unobserved factors might still influence this choice. Another reason might be that respondents have mentioned other reasons as well. To avoid this problem, another methodology can be used by taking weights into account for switching behaviour. Hereby one creates a distribution around the potential, i.e. potential is not a fixed number anymore.

For future research, a few recommendations to increase model performance can be made. A first recommendation is to refine the number of levels of the variables, e.g. building density now contains two levels: high and low, in future research this variable could contain three or more levels. A second recommendation is to include more independent variables in the models. As was seen in the literature review, the number of influencing factors is higher than the number of variables considered in this research. Again, a more refined categorization will lead to increased results. Measures and potential that are now explained by ASC's can be more nuanced in the future by including new variables and new levels. A final recommendation is to integrate the two models build here into one model, even though the models describe another dependent variable. The behaviour remains the same, namely to derive a model that defines PT potential. In a nested structure, first the model in mentioning a specific reason is considered, and then the model to determine switching behaviour is considered. Another model improvement might be to use more behaviour realistically models, e.g. a mixed logit model. Another methodology might be applied to the dataset. In this research revealed-preference data was used but it might be possible to use stated-preference data, this means however creating a whole new questionnaire. Advantageous to this method is the PT providers can create fictional situations and can derive how consumers will react.

8 BIBLIOGRAPHY

- [1] D. Janssens, M. Cools, W. Miermans, K. Declercq en G. Wets, „Onderzoek Verplaatsingsgedrag Vlaanderen 4.2 (2009-2010),” Instituut voor Mobiliteit, Diepenbeek, 2011.
- [2] Projectteam MON, „Mobiliteitsonderzoek Nederland 2009,” Ministerie van Verkeer en Waterstaat, 2010.
- [3] Y. Zhou, K. Viswanathan, Y. Popuri en K. E. Proussaloglou, „Transit District Customers in San Mateo County, California: Who, Why, Where and How,” *Transportation Research Record*, 2004.
- [4] K. J. Krizek en A. El-Geneidy, „Segmenting Preferences and Habits of Transit Users and Non-Users,” *Journal of Public Transportation*, vol. 10, nr. 3, 2007.
- [5] XTNT, „Reizigersonderzoek Openbaar Vervoer en Fiets Overijssel,” Provincie Overijssel, Zwolle, The Netherlands, 2010.
- [6] R. Buehler, „Determinants of Transport Mode Choice: A Comparison of Germany and the USA,” *Journal of Transport Geography*, pp. 644-657, 2011.
- [7] P. van der Waerden, H. Timmermans en M. Bérénos, „In search of the Public Transport Users: Towards Public Transportation Potential Maps,” *Transportation Research Board*, 2005a.
- [8] D. Dion en G. Cliquet, „Consumer Spatial Behavior,” Iste Publishing Company, 2006.
- [9] A. Banos, „From individual statements to market maps: a geomarketing approach to public transport Planning,” *Cybergeo: European Journal of Geography*, 2002.
- [10] V. Van Acker, B. Derudder en F. Witlox, „Assessing Objective and Subjective Spatial Influences of Modal Choice,” *Universiteit Gent, Gent*, 2010a.
- [11] C.-F. Chen en W.-T. Lai, „The Effects of Rational and Habitual Factors on Mode Choice Behaviours in a Motorcycle-Dependent Region: Evidence from Taiwan,” *Transport Policy*, pp. 711-718, 2011.
- [12] J. de Dios Ortúzar en L. G. Willumsen, *Modelling Transport*, Wiley, 2004.
- [13] S. Handy, „Does the Built Environment Influence Physical Activity? - Examining the Evidence,” *Transportation Research Board, Washington*, 2005.
- [14] K. Train, „Discrete Choice Methods with Simulation,” Cambridge University Press, Cambridge, UK, 2009.
- [15] V. Van Acker, B. Van Wee en F. Witlox, „When Transport Geography Meets Social Psychology: Toward a Conceptual Model of Travel Behaviour,” *UGent - Department of Geography, Gent*, 2010b.
- [16] I. Ajzen, „The Theory of Planned Behavior,” *Organizational Behavior and Human Decision Processes*, nr. 50, pp. 179-211, 1991.
- [17] P. van der Waerden, M. Bérénos en H. Tromp, „Regionaal Openbaar Vervoer Bereikbaarheid Consumentgericht,” in *Colloquium Vervoersplanologisch Speurwerk*, Antwerpen, 2009.
- [18] J. Pucher en J. L. Renne, „Urban-Rural Differences in Mobility and Mode Choice: Evidence from the 2001 NHTS,” *Rutgers University, New Brunswick, New Jersey*, 2004.
- [19] P. van der Waerden, M. Bérénos en G. Verheyen, „Validatie van een gebiedsgericht verovermiddelkeuze model,” in *Colloquium Vervoersplanologisch Speurwerk 2007*, Antwerpen, 2007.
- [20] Transit Oriented Development, April 2012. [Online]. Available: <http://www.transitorienteddevelopment.org/>. [Geopend April 2012].
- [21] H. Sung en J.-T. Oh, „Transit-oriented development in a high-density city: Identifying its association,” Elsevier, 2010.
- [22] MuConsult, „Mobiliteit eindigt bij het werk - Het effect van de werkomgeving en de woon-werkroue op de vervoerwijzekeuze,” MuConsult B.V., Amersfoort, 2001.

- [23] P. van der Waerden, M. Bérénos en R. Zonnenberg, „Gemeentelijk Mobiliteitsonderzoek (GMO): Onderzoek naar Verplaatsingsgedrag op Verschillende Gemeentelijke Schaalnivaus,” CVS-Congres, 2005b.
- [24] H. Aarts, B. Verplanken en A. van Knippenberg, „Habit and information use in travel mode choices,” Elsevier, 1997.
- [25] J. A. Ouellette en W. Wood, „Habit and Intention in Everyday Life: The Multiple Processes by Which Past Behavior Predicts Future Behavior,” *Psychological Bulletin*, vol. 124, pp. 54-74, 1998.
- [26] H. Aarts en A. Dijksterhuis, „The Automatic Activation of Goal-Directed Behavior: The Case of Travel habit,” *Journal of Environmental Psychology*, pp. 75-82, 2000.
- [27] R. Kitamura, T. Yamamoto, Y. O. Susilo en K. W. Axhausen, „How Routine is a Routine? - An Analysis of the Day-to-Day Variability in Prism Vertex Location,” Kyoto University, Kyoto, 2005.
- [28] T. Arentze, H. Timmermans, F. Hofman en N. Kalfs, „Data Needs, Data Collection, and Data Quality Requirements of Activity-Based Transport Demand Models,” TRB Transportation Research Circular, 2000.
- [29] Transumo, „Transumo-project "Regionaal openbaar vervoerbereikbaarheid consumentgericht" (ROVBECO) - Geo-onderzoek en -marketing van mobiliteit,” Transumo, Zoetermeer, The Netherlands, 2009.
- [30] SRE, 2012. [Online]. Available: www.sre.nl. [Geopend 2012].
- [31] J. de Abreu e Silva, T. F. Golob en K. G. Goulias, „Effects of Land Use Characteristics on Residence and Employment Location and Travel Behavior of Urban Adult Workers,” *Transportation Research Record*, pp. 121-131, 1977.
- [32] L. D. Frank en G. Pivo, „Impacts of Mixed Use and Density on Utilization of Three Modes of Travel: Single-Occupant Vehicle, Transit and Walking,” *Transportation Research Record*, 1994.
- [33] H. Meurs en R. Haaijer, „Spatial Structure and Mobility,” MuConsult, Amersfoort, 2001.
- [34] F. Zhao, M.-T. Li, L.-F. Chow, A. Gan en D. L. Shen, „FSUTMS Mode Choice Modelling: Factors Affecting Transit Use and Access,” National Center for Transit Research (NCTR), Tampa, 2002.
- [35] CBS, „CBS - Bevolking - Cijfers,” Centraal Bureau voor Statistiek, 2012. [Online]. Available: <http://www.cbs.nl/nl-NL/menu/themas/default.htm>. [Geopend April 2012].
- [36] D. A. Hensher, J. M. Rose en W. H. Greene, *Applied Choice Analysis - A Primer*, Cambridge, UK: Cambridge University Press, 2005.
- [37] J. van den Bergh, E. van Leeuwen, F. Oosterhuis, P. Rietveld en E. Verhoef, „Social learning by doing in sustainable transport innovations: Ex-post analysis of common factors behind successes and failures,” *Research Policy*, pp. 247-259, March 2007.
- [38] W. Veeneman en J. Koppenjan, „Securing public values in public transport projects: Four Dutch cases on innovation,” Delft University of Technology, Delft, 2010.
- [39] G. Corpuz, „Public Transport or Private Vehicle: Factors That Impact on Mode Choice,” New South Wales Ministry of Transport, Haymarket, 2007.
- [40] T. Schwanen, „Gender Differences in Chauffeuring Children among Dual-Earner Families,” *The Professional Geographer*, pp. 447-462, 2007.
- [41] D. Dissanayake en T. Morikawa, „Investigating household vehicle ownership, mode choice and trip sharing decisions using a combined revealed preference/stated preference Nested Logit model: case study in Bangkok Metropolitan Region,” *Journal of Transport Geography*, pp. 402-410, 2010.

- [42] B. Quinn, „Transit-Oriented Development: Lessons from California,” *Built Environment*, pp. 311-322, 2006.
- [43] CBS, „Wonen en werken in de stad,” CBS, 1999.
- [44] J. J. Louviere, D. A. Hensher en J. D. Swait, „Stated Choice Methods - Analysis and Application,” Cambridge University Press, Cambridge, 2000.

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
Modelling Public Transport Potential in Relation to Area Characteristics

Richting: **master in de verkeerskunde-mobiliteitsmanagement**

Jaar: **2012**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Voor akkoord,

Poesen, Wannes

Datum: **1/06/2012**