UNIVERSITY OF TWENTE.

Faculty of Engineering and Technology,
Civil Engineering and Management

Intrapersonal Variation in Destination Choice.

Saidul Chowdhury

Centre for Transport Studies

Faculty of Engineering and Technology,

Civil Engineering and Management,

University of Twente.

Enschede, The Netherlands.

Title	Intrapersonal Variation in Destination Choice.
Version	Journal paper version (V1)
Program	Transportation Engineering and Management
Date	August 25, 2017
Author	Md Saidul Azam Chowdhury Student M-CEM (s1656627) Civil Engineering & Management, University of Twente. m.s.a.chowdhury@student.utwente.nl
Supervisors	Prof. dr. ing. Karst Geurs Centre for Transport Studies University of Twente. k.t.geurs@utwente.nl Dr. ing. Lissy La Paix Puello Centre for Transport Studies University of Twente. l.c.lapaixpuello@utwente.nl
Duration	February 2017 - August 2017

UNIVERSITY OF TWENTE.

Contents

List of Figures

List of Tables

Abstract

The dynamics in travel behavior have been exploring widely in travel demand industry with a focus on interpersonal variability. And, numerous research has been conducted to analyze travel behavior. Nevertheless, intrapersonal variation of travel behavior in destination choice modeling is still underestimated. These models are capable to accommodate population heterogeneity, but still not often used in large scale. In this paper, destination choice models are developed to estimate the intrapersonal variation in travel behavior. The models incorporate the effects associated with trip characteristics and spatial information on travel behavior.

Two years (2014 and 2015) data is used to develop the model from the Dutch Mobile Mobility Panel (DMMP) by using mixed logit model. Total 68626 valid trips are recorded for 442 respondents who participated in both years. Data were collected by a smartphone app that uses global positioning system (GPS), and automatically detects departure-arrival times, origin-destinations, modes. Based on the activity purposes, data are segmented into fixed (work, education, appointment, etc.) and flexible (shopping, leisure, sports, tour, etc.) destinations. Discrete destination alternatives are defined based on individuals' behavior of destination repetition and statistical distribution of the spatial repetition index to capture the intrapersonal variation in destination choice.

The model results and probabilities show that the intrapersonal variation is high for the less repeated locations, which clearly represents the variation (or novelty) seeking behavior for choosing destinations. Nevertheless, this variation exists also in the departure time and mode repetition, which is particularly high for the less repeated locations. Elasticity revealed the connection between activity, departure time and destination. Travel time and departure time is found significant parameters. For example, people trust the bicycle and walking travel time, and while the morning is likely to travel towards fixed destinations, afternoon and evening is likely for travelling to the flexible destinations. Fixed destination trips are mostly in the commercial and industrial area and unlike to be performed during the weekend, while retail and recreation grounds are likely for flexible destinations. Accessibility of public transport is found more reliable than bicycle and car. Lastly, built environment variables are found strongly correlated with mode choice.

Keyword: *destination choice, intrapersonal variation, panel data, spatial information, mixed logit*

1 Introduction

For travel demand analyses, discrete choice technique is widely used and has been testing by the researchers. This technique can estimate and predict a decision maker's (can be a person, household, or a firm) choice of one alternative from a finite set of mutually exclusive and collectively exhaustive alternatives (Koppelman & Bhat, 2006). Most of these models are trip-based and founded on traditional four-step theory which particularly focuses on individuals travel behavior (single-choice dimensional), while activity-based demand models got attention from the last decade which is more behaviorally oriented and focuses on trip characteristics (multi-dimensional choices). Hierarchical model structures are proposed for the activity-based demand models which try to deal with multiple choices in an integrated framework (González *et al.*, 2016; Ishaq *et al.*, 2012). These models have a fixed model structure which assumed to be governed the behavior of entire sample population. However, this assumption can be incorrect, because human behavior exhibits considerable variation within and between behavioral units (R Pendyala, 1998). There can have interdependency and intrapersonal relations between choices. For instance, a group of people can choose the travel mode based on the destination choice, on the other hand, for another group, the mode choice decisions may be influenced by route choice, or car ownership, weather etc. And, it can happen that there is no influence on each other choices.

To provide efficient transport planning, understanding the variability of individual daily activity-travel behavior is necessary (Tarigan *et al.*, 2012). Although, the dynamics of travel behavior has been exploring widely in travel demand industry with a focus on interpersonal variability – behavior variation between individuals, however less focalize given to the intrapersonal variation – behavior variation within individuals. This variation can be day-to-day (e.g. Deutsch-Burgner (2015); Pas and Sundar (1995)), week-to-week (e.g. Tarigan *et al.* (2012)), or weekday-to-weekend. In addition, when there is available data for the same respondents for consecutive seasons or years, variation can be in season-to-season or year-to-year as well (Tarigan *et al.*, 2012). And this variation within individuals can be determined from the repetitive characteristics of the explanatory variables which can be trip characteristics, spatial components or may be socio-economic characteristics. The historic evidence of intrapersonal variation in travel behavior is remarkable (see Buliung *et al.* (2008)). Some recent studies explored intrapersonal variation in mode choice (Heinen & Chatterjee, 2015; Thomas & Geurs, 2016), leisure activity travel-pattern of one-and two worker households (Tarigan *et al.*, 2012), however need to be explored in more aspects of travel behavior.

Trip distribution is the prerequisite for analyzing individuals' travel behavior. In the past, gravity models were used, where trip production-and attraction are grouped and distributed based on impedance only (César A. Segovia, 2015). In gravity model, mono-centric urban areas are considered only for the trip distribution (*Transit Modeling Update*). However, the modern studies are not anymore dependent on mono-centric areas, but more about multiple and important attraction area, urban-suburban trips and interest of more sustainable development. As a result, destination choice models are widely suggested in the literature as a good replacement for gravity model (César A. Segovia, 2015). This study explored the intrapersonal variation in individuals' destination choice with the consecutive two years data from the Dutch Mobile Mobility Panel. In this regard, two types of destinations are categorized based on the flexibility of the activities for better market segmentation. Also, two types of discrete alternatives are defined for each segment based on the spatial repetition and frequency. This study will try to answer following research questions: (a) How the destinations can be defined? (b) How can the intrapersonal variation in destination choice behavior be explained? (c) What factors influence individuals' frequency of destination choice? (d) To what extent intrapersonal variation affects the choice? And (e) To what extent the model result

can be implemented? To answer these questions statistical analysis is carried out and a set of mixed logit models are developed where several trip characteristics and spatial variables are tested and influential factors are identified. Spatial variables (i.e. land use types, accessibility, built environment) are collected from different sources and joined with the DMMP to enrich the data.

This article is organized into 6 sections, while this is the section 1: introduction. Section 2 introduces the methodological literature review which consists literature about the advancement of destination choice modeling, along with the influential travel- and spatial variables. Section 3 introduces the database that is used in this study. Section 4 discusses the Methodology of this study that consists of data preparation and variable specification, methodology for defining destination definition, estimating intrapersonal variation, segmentation of the data and analytical framework for the model. Section 5 discusses the descriptive statistics, model estimation results, forecasting and elasticity. Section 6 highlights the summary of the study, discusses the possible implications of the results and provides some insight for the future research. Lastly, this report ends with references and appendices.

2 Literature Review

2.1 Destination Choice

In travel behavior analysis, destination choice is one of the crucial steps, very flexible and has the capability to accommodate population heterogeneity – varies person to person (Mishra *et al.*, 2013; Ye *et al.*, 2012). The likely existence of temporal variation in explicitly spatial dimensions (e.g., activity location, and the geography of trip-chains) of heterogeneous activity-travel decision making (Cirillo *et al.*, 2003; Schlich *et al.*, 2004; Schönfelder & Axhausen, 2010) draws attention to a potential gap in current era of travel behavior oriented demand modeling. Since destination choice models are estimated at the disaggregate level, impedance variables interact with individual's demographic and socioeconomic characteristics (Mishra *et al.*, 2013; Ye *et al.*, 2012). In this context, destination choice models can incorporate the accumulation effects related to trip chaining and spatial autocorrelation and competition (Bernardin Jr *et al.*, 2009). In addition, this type of destination choice modeling approach can estimate the effects of spatial information (e.g. land use) on travel behavior simultaneously (Chow *et al.*, 2005). As a matter of fact, a lot of studies proved the improvement of destination choice modeling, however they are not yet often used in large scale studies. Literature shows that, most destination choice studies are mainly for non-work trips (e.g. leisure, tourism, recreation), nevertheless, except few (e.g. (Newman & Bernardin, 2010)), other regular activities are underestimated in destination choice modeling. For example, for shopping or sports related activities, people have alternative choices that can depend on many factors including weather, departure time, travel time, etc. On the other hand, regular activities like work, business trips, picking up/bringing away can have some pre-fixed locations, but still several factors are responsible for the frequency of the destination repetition which is still not explored.

2.2 Influential Variables

Practically, the destination choices are very much complex and determined by several factors including the characteristics of the destination itself and the accessibility by different modes. However, mode and destination choices are usually executed simultaneously. As a result, people have a default mode-and destination choice combination for different daily activities (Buliung *et al.*, 2008; Hannes *et al.*, 2009) and they change this combination from a predefined set of alternatives, only if the default alternative is unavailable (Hannes *et al.*, 2009). This effect is even stronger for public transport users than for car drivers (Buliung *et al.*, 2008). However, the argument of mode and destination is a chicken and egg debate. For example, Hannes *et al.* (2009) stated that, for leisure and recreational trips, mode choice occurs before destination choice. In

contrast with the empirical evidence, Schüssler and Axhausen (2009) stated that mode choice does not play a significant role in destination choice modeling. Their argument was that, the decision making is carried out for the destination first and later on, the mode choice decision comes along with the consideration of the accessibility by different transport modes. Mokhtarian and Salomon (2001) also have the same indication. They stated that, for novelty seeking behavior – which is an important factor in travel behavior, particularly for the leisure and recreational activities (Arentze & Timmermans, 2005; La Paix *et al.*, 2017; Schlich *et al.*, 2004; Schönfelder & Axhausen, 2010) –, people may sometimes do excess travel, which doesn't depend on mode or route, because this novelty seeking is considered part of the attractiveness measure of the destination that came first. Moreover, sometimes decision (subconscious) to travel made first (e.g. mode choice), and then a destination or activity is invented to support that decision (Mokhtarian & Salomon, 2001). For example, the Sunday drive, which became common during the growth of automatization, a desire to see the scenic countryside and landscape that is often a destination specific motivation (Muller, 2004). Along with car, public transport, bicycle, taxi, walking can be considered as main travel mode for short distance trips for which, space-time constraints are relaxed, although weather (rain, snow, etc.) can be an external constraint, which is true for cycling as well (Hannes *et al.*, 2009). Table 1 revealed the historical evidence of explanatory variables that are influential in individuals' destination choice, although most of the studies are only for leisure and tourism.

Table 1: Influential variables in destination choice

		Explanatory variables	**Model**
Wu *et al.* (2011)	Tourism	Concretely speaking, travel time, attractiveness of destination and number of tourism spots	Latent class & nested logit
Scarpa and Thiene (2005)	Rock climbing	Mountain environment, climbers' ability, shelters, climbing routes, accessibility to site	Latent class
Simma *et al.* (2002)	Leisure (Skiing, climbing, walking, swimming)	Origin-destination distance, destination attractiveness, varied infrastructure, skiing: price level, snow condition, accessibility, weather, climbing and hiking: destination facilities (e.g. landscape, sports, cultural and eating)	Multinomial logit
Mishra *et al.* (2013)	given zone as the trip attraction end	trip char. (mode choice logsum[1], distance polynomial), household characteristics (income, auto ownership), and location characteristics (like area indicator, region-specific indicator, bridge crossing, and the size term)	Multinomial logit
Pozsgay and Bhat (2001)	recreational attraction-end choice	level-of-service (distance, price level), zonal attributes (retail and non-retail), trip attributes, and socio-demographic variables	Multinomial logit
Lepp and Gibson (2008)	Tourist role typology	Sensation seeking, gender, and an interaction term combining gender and sensation seeking	Logistic regression
Auld and Mohammadian (2011)	Non-work destination choice	Travel time, income-and race difference, land use area (e.g. recreational, residential, retail, office, institution, mixed, school, etc.), employment	Multinomial logit

Notably, several studies have been done regarding spatial and transport market. For example, Cirillo *et al.* (2003), Schönfelder and Axhausen (2003), Schlich *et al.* (2004), Buliung *et al.* (2008), Hannes *et al.* (2009), Schönfelder and Axhausen (2010), etc. Spatial information can be the geographic data related to the actual appearance of objects on the surface of the earth, such as administrative boundaries, hydrologic boundaries, land use, potential and soil characteristics, and hydrological and building irrigation networks, etc. (Falahah *et al.*, 2014). Spatial variables (e.g. land

[1] The magnitude of the logsum parameter represents the presence of common unobserved attributes affecting the attractiveness of elemental alternatives in a choice alternative (see Pozsgay and Bhat (2001))

use, accessibility, built environment) have a promising impact on travel behavior analysis, particularly in destination choice modeling. For instance, Hannes *et al.* (2009) revealed the variety of critical spatial factors in an individual's mental map that influences daily activity travel behavior. Other interesting findings on the attributes of existing routes are that cyclists have marked preferences for those that are tree-lined or include bikeway segments while tending to avoid routes used by bus lines (González *et al.*, 2016). Moreover, different type of spatial mobility constraints determines the modal variability (Heinen & Chatterjee, 2015). Although accessibility is an important factor for transport and land use planning (Handy & Niemeier, 1997), it is hardly used in practice for modeling destination choice. For non-work and tour trips, accessibility is found to be influential for destination choice (Hooper, 2015; Huang, 2014; Limanond & Niemeier, 2003). Most studies considered trip characteristics (e.g. travel time, distance, cost) as a part for accessibility measurement, whereas activity elements (spatial distribution of activities) are underused. The activity elements are the 'attractiveness' of a particular location as a trip destination (Handy & Niemeier, 1997), that can be defined by the job accessibility – no. of jobs available – in a particular location. Another important factor for urban planning is the built environment. The link between travel behavior and built environment is empirically established. For example, Chen *et al.* (2008) considered density as a measure for the built environment variable to model mode choice. González *et al.* (2016) found that, destination choice is strongly influenced by Metro stop locations, indicating that a combined bicycle-Metro mode generates a strong synergy. Therefore, built environment variable has the potential to reveal the causal relationship with destination choice and still needs to be explored.

2.3 Role of Repetition

The distribution of change in variables for time-to-time can be compared to explore the observed variation. Historically, intrapersonal variation was expressed as being the standard deviation from the average observed behavior. Since destination choice is not continuous but a frequency variable, estimating standard deviation is not applicable. This variability can be estimated in terms of different dimensions of travel behavior, such as trip purpose, mode choice, route choice, destination choice, activity duration, starting time of the activities, etc. (Heinen & Chatterjee, 2015; Schönfelder & Axhausen, 2010). For instance, Huff and Hanson (1986) assessed repetition of mode-trip purpose pair combinations. They also concluded that, individuals' travel activity patterns can be characterized by both repetition and variability. Schlich and Axhausen (2003) found a high degree of repetition in several combinations of mode, trip purpose, arrival time and destination. The evidence of repetitiveness in travel pattern is established by Stopher and Zhang (2010). However, spatiotemporal analysis can be interesting for analyzing intrapersonal variation in destination choice which is because, location based repetition varies across the type of the activity and over time (Buliung *et al.*, 2008).

Further, destination choice is dependent on the characteristics of the alternatives and of the travelers, and therefore, having information about the demand and supply side for the whole investigated area is important (Simma *et al.*, 2002). Also, there should have a pattern or similarities that can occur because of the trip purposes. Upon these similarities, different type of destinations can be defined. For instance, Schüssler and Axhausen (2009) suggested that evaluation and weighting between different type of destination choice alternative aspects and their interaction can explain better the comprehensive similarities among them. Upon the prerequisites, Schüssler and Axhausen (2009) came up with the following aspects of similarity that can be accounted for proposed the destination choice modeling framework: similarity derived from *travel mode*, similarity caused by *spatial proximity*, similarities emerging from *spatial learning and spatial repetition*, and similarities originating from the *image of the destinations*. The third approach might be the best

approach, since DMMP dataset consists multiple day observations and the evidence of location-based repetition in travel pattern (Buliung *et al.*, 2008). On the other hand, mode choice doesn't play a vital role in the most destination decision making process (as discussed earlier), although it has a strong influence on the choice set in terms of cost, accessibility, availability of mode etc. to the individual. And for this, mode similarities or mode repetition is important in destination choice modeling. Secondly, spatial proximity requires a combined route, mode and destination choice model. Lastly, the image of a location strongly depends on the spatial resolution level and the available alternative characteristics.

Not to mention, the segmentation has been used in most travel demand analyses to obtain more realistic models, which is mainly done to divide the data into homogeneous groups of individuals and to understand individuals' behavior towards the choice of alternatives (Ishaq *et al.*, 2012). This can help to predict and understand better in the choice of alternatives (Currim, 1981). According to Boyce and Bar-Gera (2004), the assumption behind segmentation is that travel-decision characteristics are the same within each class, but differ among classes. Generally, the variables that are used to segment the data, are determined according to the data analysis, the researcher's experience, and a process of trial and error (Ishaq *et al.*, 2012). However, this can be done, for example, according to the trip purpose, trip distance, SE (e.g. income group, residence location, age, gender, employment) etc. (Outwater *et al.*, 2015). To calibrate a destination choice model, this can be done by trip purpose, time period and trip market segment (as allowed by the sample size). Recently, Wang *et al.* (2016) developed a joint destination-mode choice by two market segmentation based on the trip purposes: work-related business and personal business (i.e. sightseeing, visiting friends or relatives, seeing doctors, and shopping). Therefore, segmenting the data based on trip purposes can provide better estimation result to analyze the demand and capture the intrapersonal variability for particular destination segment.

In the realm of travel behavior and discrete choice analysis, destination choice modeling has been less of a focus. Even though some researchers have explored leisure and tourism destination choice, this still needs to be explored in a larger scale (e.g. activity set, trip data). This study develops a set of destination choice model with a set of discrete destination alternatives defined from the spatial repetition and frequency and contributes to the literature. This approach is still conceptual (Buliung *et al.*, 2008; Schüssler & Axhausen, 2009) and to the best knowledge, no study used this concept so far. Moreover, since every possible trip purpose is considered, the contribution of this study to the literature will have a greater added value. The analysis is done with panel trip diary (DMMP) to explore inter-and intrapersonal variation in destination choice, associated with several socio-economic characteristics, trip characteristics and spatial variables.

3 The Dutch Mobile Mobility Panel (DMMP)

To understand travel behavior, cross-sectional survey technique is mostly used (Yáñez *et al.*, 2011). However, to understand the dynamics, intrapersonal variation and changes in travel behavior over time more accurately, this is still not sufficient. Longitudinal (or panel) data – multiple observation of the same respondent over a time-period – has the potential to overcome the above-mentioned shortcomings (RM Pendyala & Pas, 2000). Nonetheless, most of the survey now consist of one-day travel diary and few with multiple days. For instance, three-day self-completion diaries (Deutsch-Burgner, 2015; Hoogendoorn-Lanser *et al.*, 2015), 1-week activity and travel diary (Buliung *et al.*, 2008; Montini *et al.*, 2016; Stopher & Zhang, 2010), 5-week activity-travel diary (Huff & Hanson, 1986), 6-week travel survey (Bayarma *et al.*, 2007; Schlich & Axhausen, 2003). It is remarkable that DMMP is a panel dataset with travel diary of multiple weeks and multiple years, which is unique. As a result, week-to-week, year-to-year intrapersonal variation can be explored.

Not to mention, GPS data has become omnipresent in transport modeling (see Montini *et al.* (2016) for a brief history of using GPS data in travel studies) and being used nowadays (e.g. Montini *et al.* (2016); Schlich and Axhausen (2003); Schüssler and Axhausen (2009); Stopher and Zhang (2010); Zimmermann *et al.* (2017), etc.). In order to capture accurate travel behavior, appropriate data is the main prerequisite. Traditional data collection methods (e.g. paper, pen or pencil diaries) are clumsy and have a high burden on the participants. Even with the integration of computer and web-based surveys, participants are required to log their daily activities and describe the details, such as what, where, when, etc. (Deutsch-Burgner, 2015). On the other hand, GPS based data collection process is flexible and relaxes the participant burden. Further, using GPS data, researchers can identify several trip characteristics like the mode usage, stoppages, origin and destinations, etc. and that can help researchers to construct more realistic and dynamic travel diaries. Although there is a debate about the accuracy of GPS data, preliminary screening and map-matching (Montini *et al.*, 2016; Nadine Schuessler & Axhausen, 2009) help to reduce the error rate.

This study is carried out with the data from Dutch Mobile Mobility Panel (DMMP) project. Data has been collected with a smartphone app named MoveSmarter that can detect departure and arrival times, trip origin and destinations, transport modes automatically (Geurs *et al.*, 2015). The app uses GPS for the detection process and is used to collect data for consecutive 3-years for the selected respondents. Respondents were recruited from Longitudinal Internet Studies for the Social Science (LISS) panel, which is a very representative sample of the Dutch population (about 5000 households and 8000 individuals). Data has been collected for 2 weeks (April-July) in 2 batches for the first wave (2013), where 646 respondents participated and more than 25000 trips were detected. Like the first wave, second wave (2014) also collected of 2 batches data but for 4 weeks (March-July), where more than 42000 trips were detected. In the final wave (2015), also 4 weeks of data have been collected as the second wave. The quality of trip detection and mode detection (above 90%) was much better in 2015 than 2014. For the same group of respondents (who have participated in both 2014 and 2015), about 0.5 trips per person per day (pppd) extra reported in 2015 compared to 2014. Overall, higher trip rates have been detected than the Dutch national travel surveys (3.5-3.6 compared to 2.6-3.1). However, still under-registration exists (e.g. mismatch between OD location, missing of very short trips detection, battery life for long trips, forgot to take the phone etc.), otherwise true trip rates are probably higher (0.3 additional trips pppd). Battery usage is improved by some adjustments to the MoveSmarter app with especially for smartphone owners. Participants (both in 2014 and 2015) were positive about the battery consumption in 2015 than past year.

To the best knowledge, this dataset hasn't been explored much. Although, Geurs *et al.* (2015) examined the quality of automatic trip and Thomas and Geurs (2016) explored intrapersonal mode choice variation, there is a lot of room and potential for analyzing travel behavior, inter- and intrapersonal variation.

4 Methodology

4.1 Data Availability

In this paper, we make use of data of the LISS panel administered by CentERdata (Tilburg University, The Netherlands). Only 2014 and 2015 waves are selected since the data periods are the same for both year, as well as the quality of the data is better in these two waves. To understand the dynamics over time, only the respondents who stayed and participated in both year (stayers) are selected for analysis.

Several socio-economic characteristics, trip characteristics and the weather data were available in the data. In travel demand- and behavior analysis, socio-economic characteristics of the individuals play a vital role. Although in destination choice modeling, impedance variables interact with socio-economic characteristics (Mishra *et al.*, 2013; Ye *et al.*, 2012), however considered by few literature (e.g. Pozsgay and Bhat (2001)). In this study, gender, age, occupational status, educational status, marital status, urbanity level of the residential area, no. of kids in the household, net individual income, etc. are considered as socio-economic characteristics. As trip characteristics, travel time, mode choice and departure time is considered. Travel time is selected as one of the primary trip characteristics for destination choice model based on the literature (e.g. Auld and Mohammadian (2011); Wu *et al.* (2011)), although some studies considered travel distance (e.g. Pozsgay and Bhat (2001); Simma *et al.* (2002)). The role of mode choice is established in the literature for travel demand modeling, however controversial in destination choice modeling (see literature review). Departure time is considered since the temporal variation is one of the key facts for activity-travel decision making (Cirillo *et al.*, 2003; Schlich *et al.*, 2004; Schönfelder & Axhausen, 2010). Moreover, to look into spatiotemporal variation, this might be influential. The time range specification is considered as La Paix *et al.* (2017): early morning (7-9), late morning (9-12), early afternoon (12-14), late afternoon (14-17), evening (17-20), night (20-24) and midnight to dawn (24-4).

All the spatial variables are distributed over the destination postcodes. The land use information was achieved from the open street map data. The variables are spatially joined over 6-digit complete postcode area. Therefore, each postcode has set of land use attributes and thus they are taken into account as percentages for each postcode. The considered areas are allotments; cemetery; commercial & industrial; park, forest & scrub; farm, grass & orchard; meadow & vineyard; military; heath & nature reserve; recreation ground; residential; retail, etc. Since accessibility is influential variable in destination choice (Hooper, 2015; Huang, 2014; Limanond & Niemeier, 2003), this study used accessibility data – number of jobs accessible from particular postcode as the measure and the data are retrieved from the LISA dataset 2014 for jobs. Three types of accessibility measures are taken into account: public transport, bike and car. Lastly, built environment variables which turn out also an important parameter (González *et al.*, 2016) also considered for this study and the data are retrieved from the CBS (Centraal Bureau voor de Statistiek) database. Following variables are considered: distance to doctor, daycare, train station, highway onramp, supermarket, restaurant, leisure, ice skating, swimming pool.

As discussed in the literature, weather can have influence on destination choice. Therefore, the weather is considered with three levels such as clear air, cloudy and rainy weather. Days of the week are considered and categorized as La Paix *et al.* (2017): Saturday and Sunday as weekend and rests as workdays. (see Appendix B: Data availability and SelectionAvailable Variables and Sources for more details)

4.2 Destination Definition

Although most of the studies are related to leisure, recreation and tour trips, different studies defined destinations in different ways. Literature showed that, some defined destination based on spatial resolution, image of the destination, based on the activity purpose, etc. Most precise consideration is found cities as destinations (Wang *et al.*, 2016). Although, DMMP consists GPS points, destinations cannot be defined based on that, since the data was collected with optimized battery usage and therefore every second's GPS coordinates are not recorded. In addition, the app requires time to record the trip starting and ending in the background processing. Moreover, for repeated trips to the same location, the final resting GPS points can vary significantly (Schönfelder

& Axhausen, 2010). Not only the GPS accuracy but also it could happen that (car) parking can be performed in the neighborhood because of the parking availability. To avoid these issues related to the spatial accuracy of the data, 5-digit postcode[1] is a better aggregation measure in this context.

4.3 Intrapersonal Variation

To manifest the intrapersonal variation over time, the number of destinations that an individual chooses per specific time range can simply be taken into account. Therefore, this variation can be estimated by looking into the indicator variables such as trip frequency, or distance traveled to reach those destinations, or the distribution of destinations across space (Deutsch-Burgner, 2015). To analyze the intrapersonal variation in destination choice, *Spatial Repetition Index (SRI)* – developed by Buliung *et al.* (2008) is estimated based on the trip frequency of specific activity. The formulation is expressed in Eq. (1).

Since travel demand is derived from the needs and desires to participate in various activities at different times and locations (Hägerstraand, 1970), it is likely that both types of variability exist in everyday human travel patterns (Tarigan *et al.*, 2012). In this regard, *Temporal Repetition Index (TRI)* is estimated (following the formula of SRI). Such way, aggregation is measured within **weekly** activity stability by **trip purpose**. SRI and TRI are measured as the ratio of activities carried out at repeat destination postcodes and generated at repeat time ranges[2] respectively to the total number of activity destinations utilized by a respondent during a specific week. The SRI and TRI provides a measure of spatial and temporal stability in activity-travel patterns and can be expressed:

$$SRI_i = \frac{RL_i}{TA_i} \quad \text{Eq. (1)} \qquad \text{And} \qquad TRI_i = \frac{RT_i}{TA_i} \quad \text{Eq. (2)}$$

Where, TA_i is the total number of activities in the activity set, carried out by a respondent i, during the period of time t (week). RL_i is the number of repeated locations (PC5) visited and RT_i is the number of repeated time ranges departed by a respondent i over the same period for the same activity. The indexes close to 1 indicates highly repetitive spatial or temporal behavior, with most activities occurring at repeat locations or time ranges. And the indexes closer to 0 indicates less repetitive destination outcomes or less repetitive activity generation during the same time period. In other words, repetition indexes closer to 0 means high intrapersonal variability.

4.4 Segmentation of the Data

In this study, data is segmented into two classes based on individuals' trip purpose: (a) Fixed destinations that consists of home, work, business trips, personal care, education and bringing away or picking up; (b) Flexible destinations that includes Shopping, grocery, visiting, sports or hobby, hiking or sight-seeing walk, going out, free time or leisure trips and lastly other trips. Fixed destinations are those, where a respondent is bound by to visit, consisting 63% trips of the dataset. On the other hand, flexible destinations (37%) can be chosen from different alternatives. For example, the destinations for office, home, school or doctor visit are fixed in the activity space or have strong default setting (Hannes *et al.*, 2009), however destinations for shopping, recreation, tour are flexible in activity space, because, for instance, one can choose different super shop or markets, different sports activity, travel plan, etc.

[1] The 5-digit postcode (PC5) map of the Netherlands is composed of areas where the five first characters per area are the same.
[2] early morning (7-9), late morning (9-12), early afternoon (12-14), late afternoon (14-17), evening (17-20), night (20-24), midnight to dawn (24-4)

4.5 Defining the Destination Alternatives

A large number of alternatives in the universal choice set is being considered as the main challenging aspect of destination choice modeling (Auld & Mohammadian, 2011; Thill, 1992). Modeling destination choice even at postcode or municipal level can produce a large number of alternatives. Drawing subset of the alternatives from the universal choice set for each trip can deal with this situation (Simma *et al.*, 2002). Using this approach Pozsgay and Bhat (2001) defined aggregated zones as alternatives in their attraction-end model, where each zone may contain several possible elemental attraction alternatives. In this study, aggregated destination alternatives are defined based on similarities emerging from *spatial learning and spatial repetition* (Schüssler & Axhausen, 2009). Following, alternatives are designed in two ways for each segment: based on individuals' behavior of destination repetition, and based on the statistical distribution of SRI. First, alternatives are categorized into four groups based on **individuals' behavior of destination repetition (Model 1 - M1)**: Most visited (MV), Multiple visited (MuV), Equally visited (EV) and Visited once (VO). Most visited locations are defined as if a person visited a particular location (i.e. pc5) highest for a particular activity in a particular week. Afterward, there are locations that were visited multiple times (MuV) and equally (EV) for a single purpose, and remainders are visited once (VO) only within the same time frame. Table 2 shows the trip distribution with each alternative, both for the fixed and flexible destinations.

Table 2: Trip distribution among the alternatives (M1) (%)

	MV	**MuV**	**EV**	**VO**
Fixed	77	10	4	9
Flexible	28	17	21	34

For the second model, alternatives are defined based on the **statistical distribution of SRI** and literature (Model 2 - M2). The SRI distribution is different for fixed and flexible destinations. For example, for fixed destination, 25% of the sample has index up to 0.111, next 25% are within 0.112-0.308, next 25% are within 0.309-0.421 and rests are greater than 0.421. on the other hand, this distribution is 0.040, 0.041-0.056, 0.057-0.091 and greater than 0.091 respectively which is much lower than fixed destinations. With these specifications, four alternatives are defined for each destination type: Low, medium, high and very high SRI respectively.

From these two model specifications, least visited locations (e.g. EV, VO, low and medium SRI) represents the intrapersonal variety seeking behavior. Although destination alternatives are defined in two ways, however, at the end the best model type is recommended based on the goodness of fit of both models for defining the aggregated destination type and developing the discrete choice model.

4.6 Model Development

The recent advancement in discrete choice modeling is commendable and therefore widely used for the ability to measure unobservable factors. Therefore, it is possible to identify the influential variables, that effects intrapersonal variation in destination choice. ML model structure is applied in this regard by using DMMP dataset. Models are developed for two segments: (a) Fixed destinations and (b) flexible destinations. Since (a) are very much different from the (b) (from Figure 2), two models can be developed and compared. The variables are put one by one in the model. Some basic statistical tests are carried out to understand the variable characteristics that can be found in Appendix C: StatisticsStatistical Tests for Variable Selection.

4.6.1 Model Framework

Aggregated descriptive statistics is informative but provides only a limited amount of information on the variation of (e.g. week-to-week) activity-travel behavior in individual level (Deutsch-Burgner, 2015). Therefore, discrete choice modeling technique is used to verify the descriptive statistics and to provide more clear information about the intrapersonal variation and associated influential variables in the choice of destinations. The developed model is built based on the random utility maximization theory, that has been widely used to estimate discrete choice behavior (Scarpa & Thiene, 2005; Wu *et al.*, 2011). The conceptual framework is shown in Figure 1.

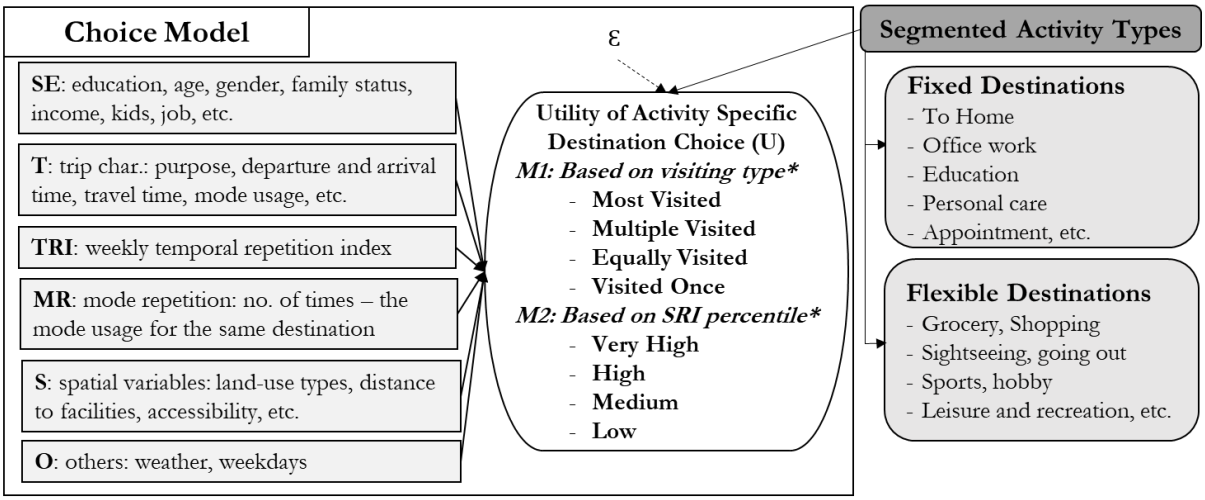

Figure 1: Conceptual Model Framework

The utility of a destination choice model is a function of multi-modal accessibilities and preferences, the attractiveness of the destination zone, person and household attributes, and other unknown, un- included attributes of the trip maker or the destination zone (César A. Segovia, 2015). The theoretical model framework is developed based on random utility theory (McFadden, 1973) as this approach offers a powerful paradigm for trip distribution modeling (Mishra *et al.*, 2013). Therefore, the structural equation is linking the deterministic model to a statistical model of human behavior.

Destination choice is choices between discrete alternatives. Mixed Logit (ML) is used to develop the destination choice model, since ML allows measuring the intrapersonal dynamics via error components, which simply create correlations among the utilities for different alternatives. The main advantage of using a mixed ordinal structure lies in having an alternative-specific setup for both respondent heterogeneity and estimated parameters. In addition, since a panel database includes repeated observations, the use of mixed multinomial logit is appropriate for panel data because it accounts for correlation among observations belonging to the same individual (Yáñez *et al.*, 2011). The utility for the ML structure can be written as (Train, 2009):

$$U_{nit} = \beta_{nit} x_{nit} + \mu_{nit} z_{nit} + \varepsilon_{nit} = \beta_{nit} x_{nit} + \omega_{nit} \qquad \text{Eq. (3)}$$

Where, *n* person has a set of alternatives *i* over choice situations *t*. U_{nit} is the utility of a destination alternative *i*. β stands for the vectors of estimated variable parameters and x_{nit} are the explanatory variables. z_{nit} is a vector of observed variable related to alternative *i* (similar like x_{nit}). μ is a vector of random terms with mean zero. ε_{nit} is the error term which is a random variable following an extreme value distribution with location parameter *0* and scale parameter *1*. The terms in z_{nit} are error components that, along with ε_{nj}, define the stochastic portion of utility. Thus, $\omega_{nit} =$

$\mu_{nit} z_{nit} + \varepsilon_{nit}$ is the unobserved utility, which is correlated with person ID and empirically depends on the specifications of z_{nj}. According to the model framework (Figure 1),

$$\beta_{nit} x_{nit} = \beta_{SE} SE_n + \beta_p T_n + \beta_{TRI} TRI_n + \beta_{MR} MR_n + \beta_s S_n + \beta_r O_n \qquad \text{Eq. (4)}$$

If φ is the vector of fixed parameters, the unconditional probability is the integral of this product over all values of ω:

$$P_{nit} = \int \overbrace{\left(\frac{e^{U_{nit}}}{\sum_j e^{U_{njt}}} \right)}^{\text{Logit probability}} \overbrace{f(\omega_{nit}|\varphi)}^{\text{Density function}} d\omega_{nit} \qquad \text{Eq. (5)}$$

To estimate the mixed logit models, simulation methods are typically used. Thus, in Eq. (5), for any given value of φ, it is possible to generate ω_{nit}^r, r=1,…,R drawn from $f(\omega|\varphi)$, which can be used later on to compute the simulated probability:

$$\check{P}_{nit} = \frac{1}{R} \sum_{r=1}^{R} \left(\frac{e^{U_{nit}(\beta_{nit} x_{nit} + \omega_{nit}^r)}}{\sum_{j=1}^{J} e^{U_{njt}(\beta_{njt} x_{nit} + \omega_{njt}^r}} \right) \qquad \text{Eq. (6)}$$

Lastly, the simulated log-likelihood (SLL) function maximizes the estimated parameters:

$$SLL(\beta_n) = \sum_n \ln(\check{P}_{nit}) \qquad \text{Eq. (7)}$$

It is interesting to see the elasticities of a given demand function with respect to changes in the values of some explanatory variables or attributes. This study explored the elasticity on the temporal dimension (with TRI). Therefore, the elasticity of a dependent variable ($\check{P}_{nit}$) with respect to another variable (TRI_i) in a function such can be expressed as follows:

$$E(\check{P}_{nit}, TRI_i) = \frac{d\check{P}_{nit}}{dTRI_i} \frac{TRI_i}{\check{P}_{nit}} \qquad \text{Eq. (8)}$$

5 Statistical Analysis

5.1 Descriptive Statistics

It is important to understand the distribution and nature of trips recorded in the dataset before conducting the analysis. For example, the number of respondents, the number of trips, trip purpose or activity types etc. As mentioned earlier, only the respondents who participated in both (2014 and 2015) wave, are considered for the study. In total 456 respondents participated, however, few of them did not complete the follow-up survey and therefore discarded. After data mining and enriching, total 442 respondents with 31441 trips in 2014 and 435 respondents with 37185 authentic trips in 2015 are recorded and selected for the analysis. Trip rates per day per person are found about 0.5 higher in 2015 (3.05) than 2014 (2.56). A general overview of the SE of the dataset is shown in the Table A 2.

5.1.1 Intrapersonal Variation

Cumulative distribution of SRI and TRI is shown below in Figure 2 which depicts that almost 60% respondents had an SRI below the median value – 0.2258, clearly represents the variety seeking behavior of the respondents. On the other hand, TRI is much lower than the SRI which explains that the variation in departure time is higher than the variation in the destination visit. Unlike TRI, SRI distribution is not uniform while in between 60-80% trips have higher repetition index in the choice of visiting destinations. The basic assumption behind segmenting is that better defining the market segments, the distribution may be more sensitive to changes in demographics, and accessibility changes (César A. Segovia, 2015). Segmentation in destination choices can reveal a large component of intrapersonal variation. Figure 2 validates the assumption, where fixed destinations have higher repetition index than flexible, in terms of both spatial and temporal index.

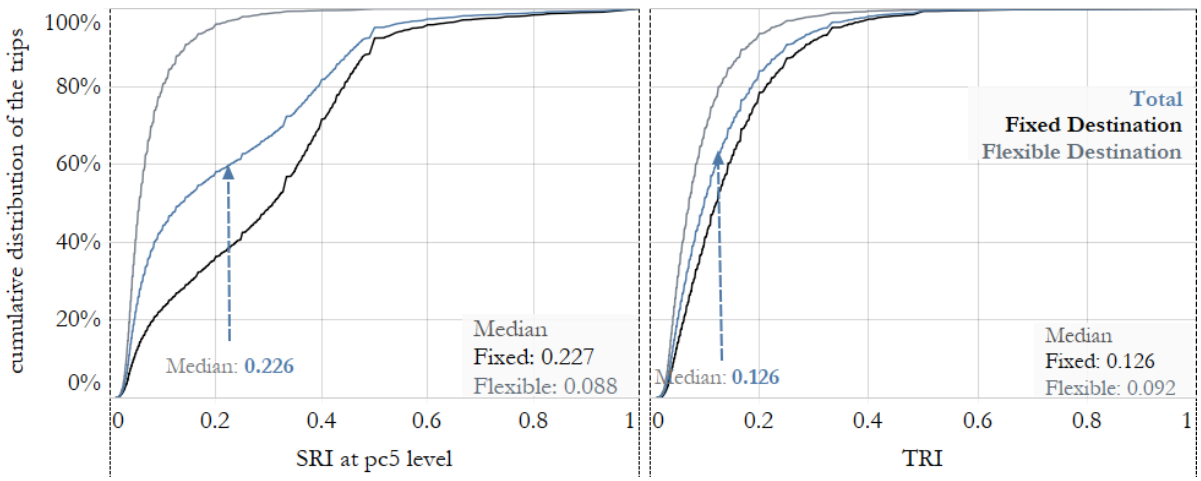

Figure 2: SRI and TRI Distribution

While more than 90% trips of flexible destination have SRI below 0.2, fixed destination trips have an index over 0.5. The same observation can be seen for the TRI as well. This context (lower value of SRI in flexible destinations) represents the variety seeking behavior of the individuals, in other words – high intrapersonal variability. In short, the intrapersonal variation in flexible destination visit is much higher than fixed destinations. To check the data quality, SRI and TRI are depicted and explained in Figure A 8 in Appendix C: Statistics Intrapersonal Variation: SRI and TRI Distribution.

5.1.2 Socio Economic Characteristics

Along with the descriptive of significant socio-economic characteristics, Table 3 explores the SRI- and TRI median[1] of fixed and flexible destinations. A glimpse is discussed in this section.

Table 3: Descriptive statistics of the SE variables in DMMP dataset for stayers (SRI, TRI – Median)

		% in sample	Trips (% by column)		Fixed		Flexible	
			Fixed	Flexible	SRI	TRI	SRI	TRI
Gender	Female	49.21	50	53	0.316	0.118	0.056	0.071
	Male	50.79	50	47	0.303	0.125	0.056	0.071
Age (years)	15-24	11.06	9	9	0.300	0.125	0.059	0.067
	25-34	14.00	14	12	0.286	0.125	0.050	0.063
	35-44	21.22	22	18	0.300	0.114	0.050	0.065
	45-54	24.38	22	23	0.278	0.125	0.056	0.074
	55-64	25.73	21	24	0.333	0.125	0.059	0.077
	>=65	15.35	10	15	0.385	0.121	0.067	0.083
Occupational status	Employed	61.44	66	54	0.280	0.125	0.050	0.065
	Retired	13.14	11	17	0.360	0.111	0.063	0.083
	Others	25.42	23	29	0.346	0.118	0.063	0.077
Education	Basic education	6.09	4	5	0.375	0.125	0.063	0.083
	VMBO	15.80	14	13	0.351	0.130	0.063	0.079
	HAVO/HBO	16.03	15	15	0.296	0.120	0.059	0.071
	MBO	25.06	27	27	0.310	0.118	0.056	0.071
	HBO	27.54	27	26	0.286	0.119	0.056	0.071
	WO	12.19	13	13	0.286	0.120	0.053	0.067
Marital status	Married	53.19	53	53	0.316	0.118	0.056	0.071
	Unmarried	28.92	30	27	0.277	0.130	0.044	0.056
	Others	17.90	17	19	0.308	0.118	0.059	0.074

[1] more robust, extreme values don't affect the median as strongly as they do the mean, useful to compare sets of data

Household composition	Single	23.25	21	24	0.310	0.118	0.059	0.071
	Married w/o child	35.21	31	34	0.333	0.126	0.059	0.077
	Married with child	39.28	40	35	0.306	0.118	0.053	0.069
	Single parent	5.87	7	6	0.250	0.111	0.046	0.063
	Others	1.81	2	2	0.257	0.141	0.048	0.065
Urbanity level of the residential area	Urban	65.7	63	66	0.304	0.125	0.056	0.074
	Semi Urban	20.9	20	19	0.333	0.121	0.059	0.071
	Not Urban	15.8	16	15	0.286	0.111	0.056	0.067

Table 3 reveals that females have more trips towards flexible destination than male and they repeat the same location more than males in fixed destinations, although TRI shows that departure time varies. The journey towards flexible destinations is same for both male and female. With age, trips towards flexible destination also increase, however, after mid-age, trip reduces towards fixed destinations. Also, in both type of destinations, both SRI and TRI increases over age. Employed respondents had the highest trip share followed by retired people. The intrapersonal variation is higher in employed group than retired people. Which means, retired people tend to visit the same place for the same purpose at the same time more than employed peoples, since the index is low for employed and higher for retired people. Trip share is higher for MBO and HBO students, but the repetition index is high for basic education and VMBO students for both type of destinations. Both the trip share and indexes are low for university students (WO) – represents that they travel less and when they do, they tend to show variety seeking behavior than habitual. Although trips towards fixed destinations can be for an appointment or educational trips. Married people have more than half (53%) trip share for both type of destinations followed by unmarried and divorced. Intrapersonal variation in flexible destination is higher for unmarried than married and other way around for the flexible. However, the repetition of departure time is lower for married people than unmarried for fixed destination and again, other way around for the flexible destinations. Respondents living in urban areas have more than 60% trips and also departure time for the trips are highly repetitive. In terms of location (SRI), semi-urban area people repeat most, followed by urban and not urban area people.

5.1.3 Trip Characteristics

In developed countries, car is now the dominant mode of transport for the people. However, in some European countries like Netherlands, Denmark; bicycle competes and has a large share in daily travel pattern. Table 4 revealed that, in DMMP dataset for the selected respondents, car is found as the dominant mode (50%) for daily travel and for all kind of activities followed by bicycle (24%) and walking (20%). It's a pity that public transport has a low share (train – 1% and bus, tram, metro – 2%). Apart from that, car is domating the trips both for fixed and flexible destination, followed by bicycle and walk (Table 4). Although, most of the modes are found havining same share percentages in both destination types, however car is found 7% high share in fixed and walking has almost double share in flexible (17%) than fixed destination trips (9%). Walking might be chosen due to the benefit of walking as a pastime or as part of a leisure activity, which is less-distance sensitive (Hannes *et al.*, 2009). Walking and cycling are sometimes considered equal alternatives in the choice set to cover short distances. Revealed decisive factors are time constraints and practical concerns; the bike is faster, but reliable storage is desired. Like walking, here too circumstances related to weather conditions and time are often mentioned as favorable preconditions. However, more often than walking, cycling is used to replace car travel for short to medium distances because of its speed and reach.(Hannes *et al.*, 2009)

Table 4 shows the median of SRI and TRI for mode choice. Taxi seems to have a high index in spatial repetition, which is unexpected. Statistics showed that, for fixed destination the share is

high for personal care (0.24%) compared to the others (i.e. to home – 0.06%, to work – 0.01% and bringing away/picking up – 0.03%). This is expected, since most elderly people use taxis in the Netherlands for going to the hospital, treatment or care center and their activity set is small in particular week. As a result, high SRI is noticed. On the other hand, going out (0.15%) and others (0.15%) has high share than shopping (0.01%), visit (0.03%) and free time (0.06%). Almost every aged people is found in this distribution. This explains, sometimes for the day out activities, people do use Taxi and if they are traveling, the activity set become small. Thus, those trips achieve a higher SRI. However, TRI represents high variation in departure time, which means above mentioned activities are least repetitive in the temporal dimension. Bicycle, moped, car is found highly repetitive in spatial dimension which is expected. Walking and other modes are also found highly repetitive. Although all modes have similar indexes, lower SRI for the train usage means high intrapersonal variability in destination choice.

Table 4: Descriptive statistics of the trip characteristics in DMMP dataset for stayers (SRI, TRI – Median)

		% in sample	Fixed	Flexible	Fixed		Flexible	
			% per segment		SRI	TRI	SRI	TRI
Mode Choice	Bicycle	24.1	23.81	24.60	0.333	0.120	0.059	0.073
	BTM	2.07	2.31	1.64	0.130	0.167	0.056	0.087
	Car	57.9	60.46	53.46	0.310	0.118	0.056	0.069
	Ferry/Boat	0.07	0.05	0.11	0.050	0.172	0.050	0.118
	Moped	1.57	1.5	1.69	0.333	0.125	0.067	0.083
	Others	0.62	0.53	0.77	0.379	0.200	0.059	0.077
	Taxi	0.05	0.05	0.04	0.405	0.080	0.077	0.069
	Train	1.93	2.44	1.04	0.091	0.184	0.044	0.097
	Walk	11.71	8.85	16.64	0.294	0.122	0.056	0.077
Travel time class (minutes)	0-7	28.91	29.04	28.71	0.333	0.111	0.061	0.071
	8-15	28.13	28.07	28.25	0.318	0.118	0.056	0.071
	16-30	24.37	24.97	23.30	0.292	0.130	0.056	0.071
	31-60	13.3	13.36	13.18	0.263	0.133	0.053	0.071
	61-120	4.41	3.83	5.39	0.267	0.118	0.056	0.071
	>120	0.89	0.72	1.17	0.257	0.103	0.056	0.071
Departure time	Early Morning (7-9)	12.54	16.56	5.55	0.152	0.152	0.063	0.059
	Late Morning (9-12)	18.53	15.65	23.53	0.263	0.095	0.059	0.083
	Early Afternoon (12-14)	14.92	12.87	18.43	0.323	0.087	0.056	0.071
	Late Afternoon (14-17)	24.47	23.18	26.72	0.345	0.139	0.056	0.083
	Evening (17-20)	18.65	18.39	19.07	0.353	0.143	0.053	0.065
	Night (20-24)	9.27	11.16	6.02	0.375	0.111	0.053	0.056
	Midnight to Dawn (24-4)	1.63	2.18	0.68	0.333	0.083	0.05	0.053

Travel time classes revealed that, 95% of the trips are performed within 1 hour and for the distribution is same for both types of destinations. Shortest travel time (0-7 minutes) has highest SRI median in both fixed and flexible destinations, which decreases with the increase in travel time. Although the share of long distance trips (>60 minutes) are low (5%), these are performed more for traveling towards flexible destinations. Unlike SRI, TRI towards flexible destinations is found same for all travel time classes. Although in fixed destinations, the direction is similar as SRI until one hour of travel time. Travel time over one hour has lower temporal repetition.

In the early morning, trips towards fixed destination are three times higher (17%) than the flexible destination (6%). Obviously work, school trips started in that time. On the other hand, late morning and early afternoon have a high share of trips for shopping, sports, leisure, etc. (flexible) than work, education, appointment, healthcare, etc. (fixed) trips (Table 4). A clearer illustration is provided in Appendix C: Statistics Figure A 2. In fixed destinations, SRI seems to be increase with

the time of the day, while lowest in the early morning, although the TRI is highest. This means, respondents start their journey for the same activity mostly in early morning (TRI- 0.152) and back in the late afternoon and evening. These are mostly work, education, appointment back home trips. On the other hand, in the flexible destination, intrapersonal variation in location choice is lowest in the early morning (since highest SRI), which means people leave early toward repeated destinations than other time periods of the day. However, TRI indicates that repetition of departure time for the same activity is higher in the late morning and late afternoon.

5.1.4 Spatial Variables

Table 5 illustrates the land use types where flexible destinations have a larger share of retail, heath and nature reserve, park, forest and scrub area than fixed destinations. On the other hand, residential, recreation ground, commercial and industrial areas have a larger trip share in fixed destinations than flexible.

Table 5: Land use and Accessibility variables (% per destination segment)

	Land use variable (percentages)											**Accessibility (no. of jobs)**		
	allotments	cemetery	commercial & industrial	park, forest & scrub	farm, grass & orchard	meadow & vineyard	military	heath & nature reserve	recreation ground	residential	retail	Bike and Ride	Car	Walk and Ride
Fixed	7	6	16	13	14	6	6	12	9	7	3	21	70	9
Flexible	6	7	8	26	11	6	6	13	2	5	10	23	64	13

Job accessibility is considered to look in depth of the destination choice. Walk and Ride (WnR), Bike and Ride (BnR) and Car stands for the number of accessible jobs by public transport, bicycle and Car. Table 5 explains very clearly that getting to the fixed destinations is much harder by relying on public transport, and much easier by car.

Furthermore, distance to different facilities from the destination, like on/off-ramp, leisure activities and attraction, restaurants, doctors, day care, etc. are considered as built environment variables. Four levels are estimated based on the percentile distribution for each variable (Table 6): low (25%), medium (25%), high (25%) and very high (25%).

Table 6: Distance to the Built Environment facilities from the destination (in km)

	Doctor	**Daycare**	**Train station**	**Highway Onramp**	**Supermarket**	**Restaurant**	**Leisure**	**Ice skating**	**Swimming pool**
Low	<0.5	<0.4	<1.3	<1	<0.4	<0.33	<2.025	<6.4	<1.5
Medium	0.5-0.6	0.4-0.5	1.3-2.3	1-1.5	0.4-0.5	0.33-0.56	2.025-4.14	6.4-14.4	1.5-2.2
High	0.7-0.9	0.6-0.84	2.4-5.4	1.6-2.2	0.6-0.9	0.57-0.92	4.15-8.2	14.5-23.4	2.3-3.6
Very High	>=1	>=0.85	>=5.5	>=2.3	>=0.9	>=0.93	>=8.3	>=23.5	>=3.7

5.1.5 Others

Presumably, most of the trips are recorded during workdays (76%) and only 24% are during the weekend (Table 7). Trips towards fixed destinations are way higher during workdays (68%), while flexible destinations are visited more during the weekend (51%). Higher SRI during weekend represents that a low intrapersonal variation in fixed location choice. For both fixed and flexible destination, higher TRI in workdays means that people repeat a particular activity within the same time range.

Table 7: Descriptive statistics of the other explanatory variables in DMMP dataset for stayers (SRI, TRI – Median)

		% in sample	**Fixed**	**Flexible**	**Fixed**		**Flexible**	
			% per segment		**SRI**	**TRI**	**SRI**	**TRI**
Weekdays (of the trip)	Workday	76	68	49	0.290	0.125	0.059	0.077
	Weekend	24	32	51	0.357	0.105	0.053	0.067

Weather condition (of the trip of the day)	clear air	36.6	37	36	0.304	0.119	0.056	0.071
	Cloudy	54.58	55	55	0.313	0.125	0.056	0.071
	Rainy	3.83	4	5	0.293	0.111	0.056	0.071
	undefined	4.59	5	4	0.333	0.133	0.067	0.077

The DMMP dataset consists of weather data where the condition of the weather, rain and temperature are available, when a certain trip is generated. More than half of the trips are found conducted during cloudy weather, followed by clear air. Since the data were collected in spring and summer time, the precipitation is found less affects the trips. Although trips towards flexible destinations were expected to have a lower share, the finding is other way around. The possible explanation can be the trips are pre-defined and thus even though it rains, respondents traveled towards a certain destination for conducting a certain activity. For fixed destinations, indexes are higher when it's cloudy, followed by clear air. However, if it rains, both SRI and TRI reduced, means people repeat the location less as well as the departure time. For flexible destination, both SRI and TRI are found uniform. The descriptive statistics and intrapersonal variability of all the explanatory variables are more elaborated and visualized in Appendix C: Statistics section.

5.2 Model Estimation

For the fixed and flexible destinations, alternatives are defined in two different ways: based on the visiting type (Model 1: M1) and based on the SRI distribution (Model: M2). For both models, alternate specific constant (ASC) are measured assuming the Most Visited (MV) location and Very high (SRI Fixed SRI>0.42, Flexible SRI >0.091) as base. Different model specifications are tested and the final models are chosen based on the informal tests (signs and magnitudes), t-test (90% confidence level) and overall goodness-of-fit measure. In this regard, pythonBIOGEME (Bierlaire, 2016) is used to estimate the models using maximum likelihood estimation (MLE) technique. To understand the panel effects, specific error components for each alternative are used in the estimated mixed logit structure. Error components are correlated with the individuals' id, so that they are creating correlation with the individuals. Models are estimated with 250 draws.

Table 8: Goodness of fit

	Model 1		**Model 2**	
	Fixed	**Flexible**	**Fixed**	**Flexible**
Number of estimated parameters:	42	55	44	43
Sample size:	42480	24425	42480	24425
Initial log-likelihood:	-5.9e+07	-3.4e+07	-6.8e+07	-3.9e+07
Final log-likelihood:	-1.6e+07	-2.3e+07	-3.8e+07	-2.7e+07
Rho-square for the initial model:	0.733	0.315	0.448	0.324

The goodness of fit (Table 8) revealed that M1 performs better than M2 – the adjusted rho square value is high in M1 than M2 for fixed destination models and other way around for the flexible destination models. Not to mention, in ML model structure socio-economic characteristics are found statistically insignificant while error components are found significant in the t-test (except the visited once alternative in M1). This is expected, since error components are creating correlation with the individuals and thus, capturing the effects of socio-economic characteristics. This is also consistent with the literature (Mishra *et al.*, 2013; Ye *et al.*, 2012).

Table 9: Model results

	Model 1: based on visiting type					**Model 2: based on SRI distribution**				
		Fixed		**Flexible**			**Fixed**		**Flexible**	
	Name	**value**	**t-test**	**Value**	**t-test**	**Name**	**value**	**t-test**	**value**	**t-test**
	ASC (Most Visited - MV)	Ref.				ASC (Very High SRI - V4)	Ref.			
	ASC (Multiple Visited - MuV)	-1.03	***-11.42***	0.335	1.47	ASC (High SRI - V3)	5.63	***41.53***	5.63	***49.38***
	ASC (Equally Visited - EV)	2.25	***11.53***	3.97	***15.93***	ASC (Medium SRI - V2)	7.81	***65.67***	5.97	***46.03***

	ASC (Visited Once - VO)	5.23	***34.8***	4.42	***38.08***	ASC (Low SRI - V1)	10.1	***61.71***	6.82	***55.17***
Mode choice	Car (MuV)			-0.164	***-3.52***					
	Train (MuV)	0.843	***8.42***			Train (V1)	0.486	***4.67***	0.445	***2.65***
	Bicycle (EV)	-0.397	***-3.95***			Moped (V3)	0.55	***4.03***		
	Bus-Tram-Metro (VO)	-0.738	***-5.24***							
	Walk (EV)			0.35	***6.32***	Walk (V2)	-0.663	***-8.77***		
Departure time	Early morning (MuV)	0.387	***7.13***							
	Late morning (MuV)	0.259	***4.91***			Late morning (V3)			-0.0695	***-1.83***
	Early afternoon (EV)			0.231	***4.16***	Early afternoon (V2)			0.0806	***1.98***
	Early afternoon (VO)			0.12	***2.41***					
	Late afternoon (EV)			0.437	***5.92***	Late afternoon (V2)	0.154	***3.87***	0.163	***3.04***
	Late afternoon (VO)			0.266	***3.98***					
	Evening (VO)			0.215	***2.24***					
	Evening (EV)			0.45	***4.33***					
Travel time (minutes)	Car (MuV)	-2.86e-05	***-5.04***			Car (V2)	-8.98e-06	***-1.9***		
	Car (EV)			2.82e-05	***4.16***	Car (V1)	-4.04e-05	***-6.69***		
	Car (VO)			2.47e-05	***4.23***					
	Bicycle (MuV)	2.34e-05	***4.71***			Bicycle (V3)			2.09e-05	***3.99***
	Bicycle (EV)			1.74e-05	***3.12***	Bicycle (V2)			1.89e-05	***3.34***
	Bicycle (VO)	3.73e-05	***6.27***	1.91e-05	***3.79***	Bicycle (V1)	3.08e-05	***7.49***	3.43e-05	***5.82***
	Public Transport (MuV)	-2.44e-05	***-4.99***			Public Transport (V3)	-2.33e-05	***-6.53***	-2.2e-05	***-4.76***
	Public Transport (EV)			-2.45e-05	***-4.52***	Public Transport (V2)	-4.68e-05	***-11.68***	-2.3e-05	***-4.63***
	Public Transport (VO)	-3.54e-05	***-6.27***	-1.79e-05	***-3.72***	Public Transport (V1)	-7.10e-05	***-14.05***	-3.4e-05	***-6.45***
	Walk (EV)	8.29e-05	***6.58***			Walk (V3)	3.03e-05	***4.3***		
	Other (MuV)			-8.06e-05	***-3.75***	Walk (V1)	9.63e-05	***11.52***		
	Others (EV)			-5.65e-05	***-2.58***					
Repetition index	Temporal repetition index (MuV)	1.59	***7.5***	3.97	***11.57***	TRI (V3)	-5.44	***-29.33***	-8.75	***-25.57***
	Temporal repetition index (EV)	-11	***-15.36***	-1.16	***-2.61***	TRI (V2)	-8.84	***-42.76***	-14.8	***-33.89***
	Temporal repetition index (VO)	2.4	***7.22***	3.67	***10.01***	TRI (V1)	-15.4	***-52.53***	-23.5	***-43***
	Mode repetition (MuV)	-0.211	***-49.77***	-0.313	***-26.96***	Mode repetition (V3)	-0.00598	***-4.06***	-0.239	***-26.74***
	Mode repetition (EV)	-2.09	***-27.97***	-2.63	***-43.96***	Mode repetition (V2)	-0.0622	***-39.41***	-0.462	***-33.28***
	Mode repetition (VO)	-4.19	***-38.35***	-3.34	***-54.33***	Mode repetition (V1)	-0.365	***-76.23***	-0.646	***-40.51***
Land use (percentages)	Recreation ground (MuV)			0.54	1.34	Recreation ground (V3)			0.484	1.48
	Park, forest, scrub (MuV)			-0.402	***-3.45***	Park, forest, scrub (V3)			-0.361	***-3.54***
	Park, forest, scrub (EV)			-0.0264	***-1.76***	Park, forest, scrub (V2)			-0.367	***-4.86***
	Park, forest, scrub (VO)			-0.448	***-7.61***	Park, forest, scrub (V1)			-0.246	***-3.57***
	Retail (MuV)			0.684	***3.52***	Retail (V3)			0.6	***2.92***
	Retail (VO)			0.738	***5.37***	Retail (V2)			0.659	***3.33***
						Retail (V1)			0.74	***3.59***
	Commercial, industrial (EV)	0.269	***6.37***			Commercial, industrial (V3)	2.12	***13.63***		
	Commercial, industrial (VO)	0.111	***3.14***	0.284	***3.77***	Commercial, industrial (V2)	3.88	***24.6***	0.207	***2.68***
	Residential (MuV)	-0.378	***-11.46***			Commercial, industrial (V1)	3.88	***24.4***		
	Residential (EV)	-0.534	***-6.29***			Residential (V2)	-0.911	***-16.53***		
	Residential (VO)	-0.422	***-6.3***			Residential (V1)	-1.15	***-18.2***		
Accessibility	Walk and ride (MuV)	3.59e-06	***11.35***			Walk and Ride (V2)	7.37e-06	***17.43***		
	Walk and ride (EV)			8.91e-07	***2.04***	Walk and Ride (V1)	1.03e-05	***19.95***		
	Walk and ride (VO)	1.97e-06	***4.84***	8.49e-07	***1.95***	Bike and Ride (V3)	1.47e-06	***5.8***		
	Car (VO)			-3.03e-07	***-2.44***	Car (V2)			3.11e-07	***3.1***
						Car (V1)	-5.71e-07	***-3.91***		
Built environment (avg. distance to the facility)	Restaurant-high (MuV)			0.153	***2.45***					
	Restaurant-medium (MuV)			0.198	***2.95***					
	Restaurant-low (MuV)			0.226	***3.28***	Restaurant-low (V2)			0.289	***4.8***
	Restaurant-low (EV)			0.239	***3.15***	Restaurant-low (V1)			0.336	***4.98***
	Restaurant-medium (EV)			0.145	***2.13***					
	Restaurant-medium (VO)			0.128	***2.07***					
	Restaurant-low (VO)			0.175	***2.65***	Leisure-low (V1)			0.104	1.62
	Leisure-low (EV)			0.0738	0.95	Leisure-high (V2)			-0.0822	***-1.89***
	Leisure-medium (EV)			0.0708	1.35	Ice skating-high (V3)			-0.132	***-2.79***
	Ice skating-medium (VO)			0.0886	***1.91***	Doctor-high (V3)	-0.242	***-3.83***		
	Swimming pool-medium (EV)			-0.183	***-3.52***	Doctor-high (V2)	-0.527	***-8.87***		
	Swimming pool-low (EV)			-0.154	***-3.12***	Doctor-high (V1)	-0.352	***-5.11***		
	Doctor-high (MuV)	0.167	***2.93***			Daycare-high (V2)	-0.0982	***-2.44***		
	Doctor-low (EV)	0.223	***2.62***			Train station-medium(V2)	0.221	***5.92***		
	Daycare-high (MuV)	0.291	***4.99***			Train station-high (V1)	-0.255	***-5.19***	-0.139	***-3.02***
	Daycare-medium (MuV)	0.142	***2.62***			Train station-low (V3)			0.252	***4.35***
	Train station-high (MuV)	-0.207	***-3.68***			Train station-medium(V3)			0.0566	1.34
	Train station-high (VO)	-0.214	***-3.26***	-0.044	-1.03	Highway onramp-low (V3)			0.099	***2.14***
	Highway onramp-medium(MuV)	0.12	***2.26***			Highway onramp-low(V2)			0.142	***2.9***

	Highway onramp -high (VO)			0.091	***2.25***	Highway onramp-high(V2)	0.168	***4.59***		
Others	Weekend (MuV)	-0.5	***-8.54***			Weekend (V2)	-0.668	***-17.55***	0.0599	***1.66***
	Weekend (EV)	-0.492	***-4.92***	0.127	***3.25***	Weekend (V1)	-1.04	***-19.86***		
	Weekend (VO)	-0.332	***-4.25***							
	Rain (VO)			0.139	***1.7***	Rain (V2&1)	0.121	1.63		
Error comp.	σ_{MV}	-1.49	***-30.78***	-0.825	***-20.64***	σ_{v4}	-4.85	***-55.03***	4.15	***44.39***
	σ_{MuV}	-1.12	***-22.57***	-0.625	***-15.82***	σ_{v3}	1.18	***38.68***	-0.844	***-20.96***
	σ_{EV}	1.68	***20.68***	1.19	***24.69***	σ_{v2}	-0.512	***-19.69***	0.954	***14.31***
	σ_{VO}	-0.00509	-0.06	0.084	1.03	σ_{v1}	0.906	***20.77***	2.04	***38.62***

5.2.1 Model 1: Based on visiting type (M1)

The results are shown in Table 9. The result shows that, ASC of VO locations has the highest value among the others, followed by EV and for both types of destinations. This means, if rest remains constant VO locations are the most preferred travel destinations, followed by the EV and MuV. Therefore, it seems respondents are likely to explore new locations, which is consistent with the literature (Arentze & Timmermans, 2005; La Paix *et al.*, 2017; Schlich *et al.*, 2004; Schönfelder & Axhausen, 2010).

Trip characteristics

Mode usage, departure time, travel time for the chosen mode are tested in the model as alternative specific. The coefficient of mode choice is found insignificant most of the case. This is expected since Schüssler and Axhausen (2009) claimed that, mode choice does not affect much in destination choice. Moreover, built environment variables interact with most trip characteristics. For example, the excluded trip variables have a correlation with built environment variables most of the time (Table A 9). However, findings showed that, for repeated fixed locations like office, home, work, hospital, education, etc., people are likely to use the train, while bicycle and bus-tram-metro are unlikely to be used for less visited (EV and VO) same locations. On the other hand, multiple visited flexible destinations (e.g. shopping, sports, leisure and sports) are found not likely to be done by car, although for less repeated locations, walking is preferred. As expected, morning is found positively associated with multiple visited fixed destinations. On the other hand, people are likely to travel towards less repeated (EV and VO) flexible destinations during whole afternoon and evening. Further, results show that mode specific travel time is an important factor. Private vehicle (Car and bicycle) travel time is found positively associated and highly significant in equally- and only visited locations, which means people trust the travel time of private transport while exploring. The reliability of the public transport travel time is just the opposite of private transport. On the other hand, travel time of car and public transport is not trustworthy towards multiple visited fixed locations as they are significant and negatively associated, however, travel time of bicycle is positive. Multiple visited fixed locations are mostly work, education and home trips and it may happen often that traffic congestion or uncertain maintenance or accident can cause a delay in car and public transport, while bicycle overcome such issue, although may be for shorter travel time because of its speed and reach (Hannes *et al.*, 2009).

The mode repetition and TRI are added in the model as adjustment variables and are found statistically significant in all alternatives and same indication for both fixed- and flexible destinations. With respect to the most visited locations, mode repetition is highly unlikely to occur. This is expected since the variable is defined by the activity, thus lower the activity repetition, more unlikely to repeat the mode usage. In other words, mode repetition is high in most visited locations. This is consistent with Buliung *et al.* (2008). By examining location-based repetition in travel patterns, they concluded that people don't only conduct a big part of their activities at repeated locations but that they also visit these locations very often using the same transport mode. Result revealed that there is a strong link between temporal- and spatial variation (Cirillo *et al.*, 2003;

Schlich *et al.*, 2004; Schönfelder & Axhausen, 2010). For instance, TRI is found positively associated with the multiple- and only visited locations and negatively with the equally visited locations. Thus, although people repeat same time range to multiple visited locations, the tendency to repeat the departure time is high also for travelling to different places.

Spatial Variables

Several land use type variables, accessibility measures and built environment variables are tested. As expected, people are more likely to travel multiple times to a certain recreational ground and retail area as flexible destinations. also for the only visited locations, retail area is likely to be visited, which means people are likely to explore new retail areas, may be an outlet. Park, forest and scrub area are found negatively associated with all alternatives and thus, highly unlikely to be visited, which means these areas are likely to be visited as most visited flexible locations. On the other hand, commercial and industrial areas are likely to be visited as less repeated fixed locations (EV and VO). Since, residential areas are found negatively linked with the alternatives, means these areas are more likely to be visited as most repeated (MV) locations instead of others.

Since BnR (Bike and Ride) and WnR (Walk and Ride) are highly correlated, it is expected that one of them will cancel another's effect. Consequently, WnR is found significant over BnR and for less visited flexible locations (EV and VO). Relying on the car to reach only visited locations for shopping, sports, tour seems very unlikely. On the other hand, increased accessibility of public transport (WnR) can increase the utility of multiple-and only visited fixed locations

Distance to a certain facility from the destination postcode is used as the built environment variables. Results are expected, for example, distance to the restaurants are found for flexible destinations. However, for less visited locations (EQ and VO), people are more likely to prefer this distance to be within low and medium, as well as the distance to the leisure facilities. Ice skating and swimming is very popular sort of sports in the Netherlands, however model shows that only visited locations are likely for ice skating if the distance is medium and equally visited locations are unlikely to travel if the distance to swimming is low or medium. Lastly, only visited locations are found highly likely to travel even if the distance to onramp is high, while unlikely if the nearest train station is far away from the destination. on the other hand, different attributes are tested for fixed destinations. High distance to the doctor and daycare is found statistically significant for multiple visited locations. For the same type of locations, people do not prefer high distance from to the train station, however they prefer medium distance to the highway onramp.

Others

Weekends are found not likely to generate trips to any of the alternatives of the fixed destinations with respect to the most visited locations. On the other hand, flexible destinations (EV) are likely to be travelled. This is expected since work, appointment, etc. trips are mostly conducted during workdays and the trips for travelling, tourism, etc. are more likely to be performed during the weekend. When people travel to the only visited locations, rain affects positively. These trips may be determined by the activity, for instance, if someone plans to visit somewhere, even though it's raining, they perform the trip towards these locations.

5.2.2 Model 2: Based on the statistical distribution of SRI (M2)

In the SRI distributed alternatives, ASC shows that Low SRI (=<0.040) is most preferred for flexible destinations followed by medium- (0.041-0.056) and high SRI (0.057-0.091). Like model 1, this also represents the variety seeking behavior and additionally, higher intrapersonal variation in destination choice (than M1). On the other hand, medium SRI (0.112-0.308) is highly preferred

for fixed destinations followed by the low- (=<0.111) and high SRI (0.309-0.421). This is expected since most office, education, appointment and home trips repeat in the same location.

Trip Characteristics

The performance of mode usage variable is even worse than the other model (M1). This is expected as discussed in the literature (Schüssler & Axhausen, 2009). Not to mention, train is found highly preferred for traveling to the destinations with low SRI (for both fixed and flexible). Usually, long, business trips and outing trips have lower repetition index and train seems to be the preferred mode for such travel. For high SRI fixed destinations, moped is likely to be used and walking is not preferred for medium SRI destinations. Late morning is found unlikely to depart towards high repetitive flexible destinations (high SRI), while whole afternoon period seems very much likely to generate the trips for the medium repetitive destinations (medium SRI). Like the other model, bicycle travel time is found trustworthy for all the alternatives of shopping, leisure and tour activities and only low SRI alternative in fixed destination model. For both fixed- and flexible destinations, public transport travel time is unreliable, and for all the alternatives. In fixed destinations, car travel time is also negatively linked with medium- and low SRI destinations. On the other hand, walking travel time is positively associated with high and low SRI alternatives.

The performance of mode repetition variable is same as the model 1 and therefore convey the same message, however TRI is different. TRI is negatively associated with all alternatives which mean an increase in TRI (repeated departure time for the same activity), decreases the utility of SRI alternatives. This clearly represents the intrapersonal temporal variation in spatial dimension (Cirillo *et al.*, 2003; Schlich *et al.*, 2004; Schönfelder & Axhausen, 2010).

Spatial Variables

For both fixed- and flexible destinations, the performance of most spatial variables (land use, accessibility, built environment) is similar to the Model 1. For example, in flexible destination model, recreation ground area is likely as high SRI alternative, park, forest and scrub areas are unlikely for all alternatives, retail areas are likely to be visited (all alternatives). on the other hand, in fixed destination model, commercial and industrial areas are positively linked with all the alternatives, residential areas are negatively linked with medium and low SRI defined alternatives.

Unlike model 1, only car accessibility is found influential for the flexible destination – positively associated with the alternative with medium SRI. On the other hand, BnR (high SRI) and WnR (medium- and low SRI) are found positively associated with fixed destination alternatives. Therefore, relying on the bicycle to reach high SRI defined locations and public transport (WnR) to reach medium- and low SRI defined locations for work, education, home, appointment, etc. are very likely. Like the other model, car reliability to reach low SRI defined locations are found unlikely.

Further, distance to restaurant, leisure, ice skating, swimming pool, etc. are found influential for flexible destinations. For instance, if the distance to a restaurant from the destination is low, medium- and low SRI alternatives are likely to be performed; if the distance to leisure is high, medium SRI alternative is unlikely to be performed; also for the ice skating- if the distance is high, high SRI alternative is unlikely. Distance to doctor (all alternatives) and daycare (medium SRI) affect the fixed destination utility. If the distance is high choice sets are unlikely to be chosen. Although, medium distance from the train station to fixed- (medium SRI) and flexible (high SRI) destinations is influential, however high distance is not acceptable (low SRI), which is expected.

Others

Weather condition has a similar effect as for the other model (M1). As a result, weekends are unlikely to travel towards fixed destination (medium- and low SRI alternatives), while likely for flexible destinations (medium SRI alternative). Unlike model 1, rain is found positively influential for medium and low SRI defined fixed destinations. This is expected since most work or appointment trips can be carried out even though it's raining.

5.3 Implication: Elasticity and Forecasting

The elasticity explains how powerful is the temporal repetition index (TRI) to predict intrapersonal variation in destination choice. Estimated elasticities (with TRI, see Eq. (8)) and probabilities for the both models (Model 1 and Model 2) for flexible destinations are presented below in Figure 3:

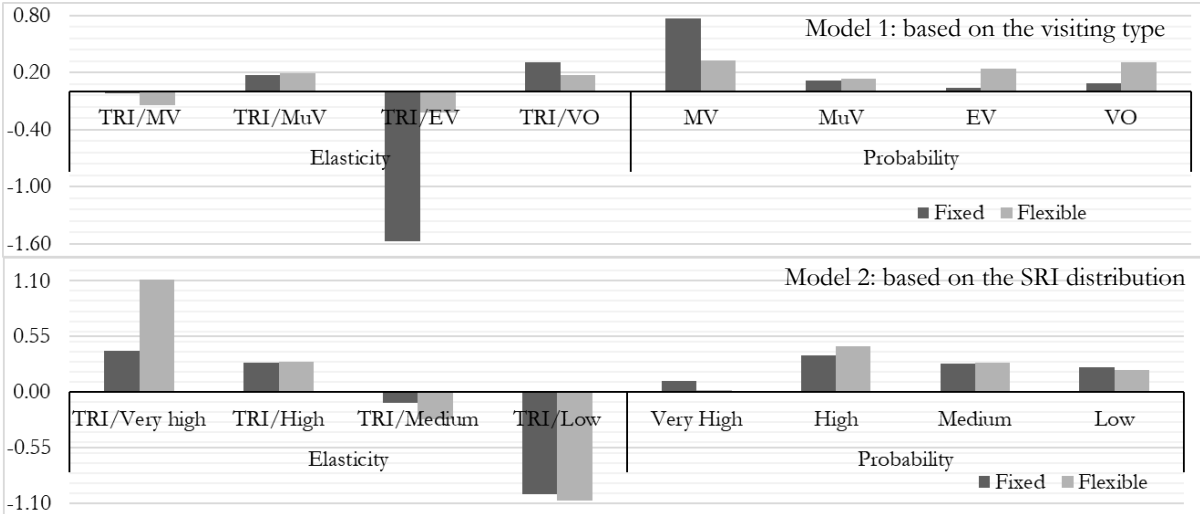

Figure 3: Elasticity and Probability from Model 1 and Model 2

The elasticity of Model 1 shows that, the demand for only visited and multiple visited location increases if the departure time of the same activity repeated more (increase in TRI). On the other hand, in Model 2, the demand for very high and high SRI-distributed alternatives seems to be increased with the increase in TRI, while other way around for the less repeated alternatives (medium and low). This means the departure time varies a lot more in the less repeated destinations. Also, in Model 1, TRI is found has highest elasticity for equally visited (EV) fixed- (-1.58) and flexible (-0.22) destinations. This means that the TRI is very important for work, home, education, appointment, etc. trips (fixed), as well as leisure, sports, shopping, etc. (flexible) and TRI is strongly linked with the spatial repetition. The same can be seen in the low SRI distributed alternative in Model 2. Since the alternatives in two models are not equally balanced, it could happen that MV (Model 1) intersects with very high or high (Model 2), but it is unknown up to which point, MV contains some of the MuV in some cases. Similarly, VO (Model 1) might be similar to low (Model 2), but EV can intersect with low or medium at the same time. In such case, perhaps EV and low are showing the similar meaning. In addition, elasticity also shows that people have so different patterns (Model 1) and also so similar (Model 2) within the same person, however Model 1 represents better the differences than Model 2. Moreover, while analyzing individuals' destination choice and classifying by SRI, Model 2 shows almost symmetric situations (e.g. flexible destination). However, while classifying by individuals' pattern (more adjusted to each person), Model 1 is more relative. Further, elasticity clearly shows that there is a connection between what people do, where and when. So, it means the distribution of activities has peak hours as well. Therefore, adjusting transport strategies to temporal and spatial distribution is necessary.

As expected, the probability of traveling to most visited locations is higher in both models. The high probability of only visited flexible locations (Model 1) clearly represents the variety seeking intrapersonal behavior for leisure, tourism, shopping, sports, etc. activities. The probability of visiting high, medium and low repeated locations are found much higher than very high repeated locations (Model 2). Clearly, also validates the variety seeking behavior in the selection of destination.

5.3.1 Additional Estimation on Elasticity and Forecasting

More than 80% trips are found having shorter travel time (up to 30 minutes). Since the intrapersonal variation is high in flexible destinations, this study estimated four different models (Model 1 and Model 2: 1-15 minutes and 16-30 minutes) to estimate the elasticity and probability for shorter travel time flexible destination trips.

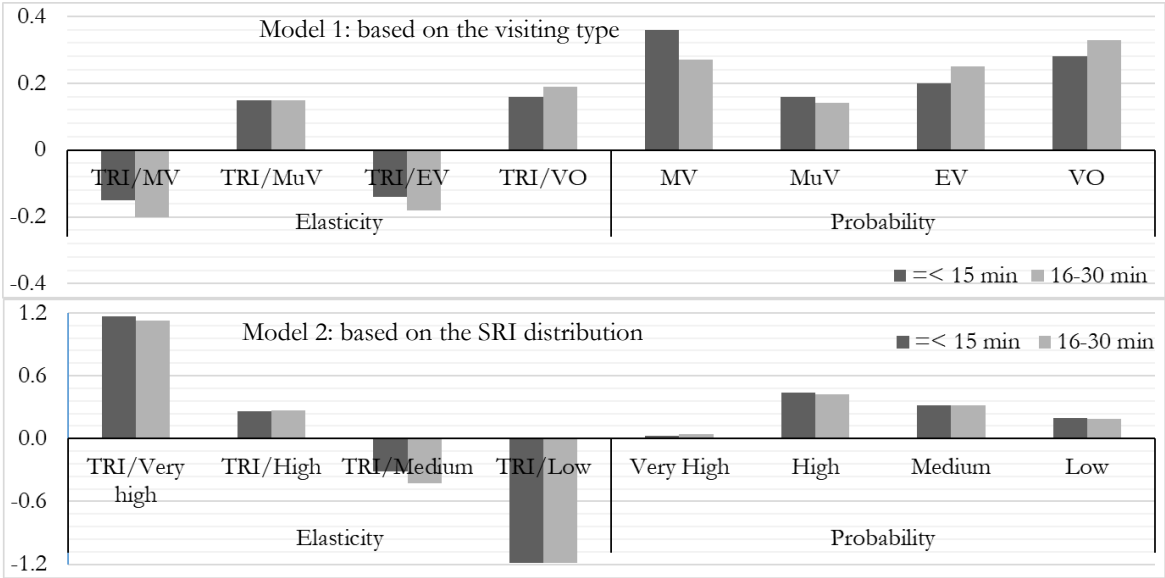

Figure 4: Elasticity and Forecasting for Shorter trips towards flexible destinations

Figure 4 shows the similar pattern as Figure 3. Model 1 shows that, elasticity of MuV, EV and VO seems to have a higher value for 16-30 minutes trips. This means, for these alternatives TRI is important. On the other hand, elasticity of Model 2 is almost symmetric. The probability in Model 2 shows that, intrapersonal variation seeking behavior is high for 16-30 minutes trips since EV and VO have a higher value. So, the probability of most- and multiple visited locations are higher for up to 15 minutes travel time. In model 2, the probability of high, medium and low SRI is high. Again, Model 1 explains the differences better than Model 2.

6 Conclusion

Although destination choice models are flexible and can capture population heterogeneity, however very much complex in practice since several factors are responsible for the choice. A set of mixed logit models are developed based on the discrete choice technique to explore those influential factors. This paper explored inter- and intrapersonal variation over time (week-to-week). Specific error components are considered for the model to account for heterogeneity effects between individuals. The destinations are defined based on the flexibility of the activity type: fixed and flexible. For each type of destinations, choice set is defined based on the spatial learning and spatial repetition (Schüssler & Axhausen, 2009). To analyze the intrapersonal variation in spatiotemporal dimension, spatial repetition index and temporal repetition index are estimated. In this regard, 5-digit postcodes are considered because of the spatial accuracy issue of the data. While most behavior analysis based transport policy is using short-term travel survey (e.g. 1, 2, 3 days or

1-week), this study used DMMP dataset that consists multiple weeks of multiple year records for the same person. Therefore, the estimated measures are more trustworthy and efficient for long-run policy and travel dynamics.

To develop the model, several socio-economic characteristics, trip characteristics, spatial variables, and other variables are considered. In the mixed logit structure, error components are found significantly interact with the socio-economic characteristics. The model result and probability reveals the existence of the variation seeking behavior among the individuals in destination choice. Mode choice is found not very much influential, which means the decision for the destination is carried out first (Mokhtarian & Salomon, 2001; Schüssler & Axhausen, 2009). However, the travel time of particular mode is found plays an important role in destination choice. An interesting finding is the existence of temporal- and modal variation in the decision-making of destination of the same activity. For instance, departure time varies a lot as well as the mode usage for traveling towards the less repeated locations. Moreover, the result of the elasticity shows a strong temporal dynamic in mobility. Spatial variables (land use, accessibility, built environment) are found associated with mode choice and explained the causal relation with the destination alternatives.

The results of this study can provide significant insights to the corresponding authorities for better transport planning and policy making/redesigning, which can lead a (long-run) change in travel behavior among the people. For example, destination characteristics (e.g. supermarket, recreational center, etc.) or built environment (e.g. infrastructure) can be redesigned/rescheduled. Also, the improvement of accessibility towards destinations by means of different transport. This study revealed the link between activity type, destination and departure time. Therefore, transport measures should be adjusted with spatiotemporal distribution and cannot be flat. So, for example, adjustment with highway and public transport, adjustment with train schedule and price with peak hours, even toll prices can be adjusted with the peak time, etc. Results are more reliable than traditional travel diary analysis in traffic planning because of the consecutive multi-day travel data, to some extent consecutive multi-week and multi-year. Furthermore, as Tarigan *et al.* (2012) stated, the results of variability's measures may be useful to capture behavioral responses and market trends from the public related to given mobility management policies (e.g. subsided public transport) over a span of time. This information is essential for policy planning and other improvements in order to guarantee the transportation system have a consistently good quality.

In future work, this study can extend to analyze day-to-day, year-to-year intrapersonal variability. Also, a combined destination-mode choice can be developed based on the repetition index that can reveal more interesting factors that affect the travel behavior of individuals.

Acknowledgements

This thesis is submitted to the Centre for Transport Studies, University of Twente in partial fulfillment of the requirements for the MSc degree in Transport Engineering and Management. This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Thanks to the Mobidot, NWO and CentERdata for the collaboration in the development of the Dutch Mobile Mobility Panel. Special thanks to John Pritchard for his support and providing the accessibility data.

References

Arentze, T. A., & Timmermans, H. J. P. (2005). Information gain, novelty seeking and travel: a model of dynamic activity-travel behavior under conditions of uncertainty. *Transportation Research Part A: Policy and Practice, 39*(2–3), 125-145. doi:http://dx.doi.org/10.1016/j.tra.2004.08.002

Auld, J., & Mohammadian, A. (2011). Planning-constrained destination choice in activity-based model: agent-based dynamic activity planning and travel scheduling. *Transportation Research Record: Journal of the Transportation Research Board*(2254), 170-179.

Bayarma, A., Kitamura, R., & Susilo, Y. (2007). Recurrence of daily travel patterns: stochastic process approach to multiday travel behavior. *Transportation Research Record: Journal of the Transportation Research Board*(2021), 55-63.

Bernardin Jr, V., Koppelman, F., & Boyce, D. (2009). Enhanced destination choice models incorporating agglomeration related to trip chaining while controlling for spatial competition. *Transportation Research Record: Journal of the Transportation Research Board*(2132), 143-151.

Bierlaire, M. (2016). PythonBiogeme: a short introduction. *Report TRANSP-OR 160706, Series on Biogeme. Transport and Mobility Laboratory, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne, Switzerland.*

Boyce, D., & Bar-Gera, H. (2004). Multiclass combined models for urban travel forecasting. *Networks and Spatial Economics, 4*(1), 115-124.

Buliung, R. N., Roorda, M. J., & Remmel, T. K. (2008). Exploring spatial variety in patterns of activity-travel behaviour: initial results from the Toronto Travel-Activity Panel Survey (TTAPS). *Transportation, 35*(6), 697. doi:10.1007/s11116-008-9178-4

César A. Segovia. (2015). Destination choice modeling primer. AECOM.

Chen, C., Gong, H., & Paaswell, R. (2008). Role of the built environment on mode choice decisions: additional evidence on the impact of density. *Transportation, 35*(3), 285-299. doi:10.1007/s11116-007-9153-5

Chow, L.-F., Zhao, F., Li, M.-T., & Li, S.-C. (2005). Development and evaluation of aggregate destination choice models for trip distribution in Florida. *Transportation Research Record: Journal of the Transportation Research Board*(1931), 18-27.

Cirillo, C., Cornélis, E., Legrain, L., & Toint, P. (2003). *Combining spatial and temporal dimensions in destination choice models.* Paper presented at the European Transport Conference, Strasbourg, France.

Cody, R. (2008). Data cleaning 101. *Robert Wood Johnson Medical School, Piscataway, NJ.*

Currim, I. S. (1981). Using Segmentation Approaches for Better Prediction and Understanding from Consumer Mode Choice Models. *Journal of Marketing Research, 18*(3), 301-309. doi:10.2307/3150971

Deutsch-Burgner, K. (2015). Multiday Variation in Time Use and Destination Choice in the Bay Area Using the California Household Travel Survey *Multiday GPS Travel Behavior Data for Travel Analysis* (pp. 34-116): Federal Highway Administration.

Falahah, Rosmala, D., & Dolok, A. M. (2014). *Socialization and Visualization of City Transport Using Google Maps API.*

Geurs, K. T., Thomas, T., Bijlsma, M., & Douhou, S. (2015). Automatic Trip and Mode Detection with Move Smarter: First Results from the Dutch Mobile Mobility Panel. *Transportation Research Procedia, 11*, 247-262. doi:http://dx.doi.org/10.1016/j.trpro.2015.12.022

González, F., Melo-Riquelme, C., & de Grange, L. (2016). A combined destination and route choice model for a bicycle sharing system. *Transportation, 43*(3), 407-423. doi:10.1007/s11116-015-9581-6

Hägerstraand, T. (1970). What about people in regional science? *Papers in regional science, 24*(1), 7-24.

Handy, S. L., & Niemeier, D. A. (1997). Measuring Accessibility: An Exploration of Issues and Alternatives. *Environment and Planning A, 29*(7), 1175-1194. doi:doi:10.1068/a291175

Hannes, E., Janssens, D., & Wets, G. (2009). Does space matter? Travel mode scripts in daily activity travel. *Environment and Behavior, 41*(1), 75-100.

Heinen, E., & Chatterjee, K. (2015). The same mode again? An exploration of mode choice variability in Great Britain using the National Travel Survey. *Transportation Research Part A: Policy and Practice, 78*, 266-282. doi:http://dx.doi.org/10.1016/j.tra.2015.05.015

Hoogendoorn-Lanser, S., Schaap, N. T. W., & OldeKalter, M.-J. (2015). The Netherlands Mobility Panel: An Innovative Design Approach for Web-based Longitudinal Travel Data Collection. *Transportation Research Procedia, 11*, 311-329. doi:http://dx.doi.org/10.1016/j.trpro.2015.12.027

Hooper, J. (2015). A destination too far? Modelling destination accessibility and distance decay in tourism. *GeoJournal, 80*(1), 33-46. doi:10.1007/s10708-014-9536-z

Huang, Y. (2014). *Accessibility and non-work destination choice: A microscopic analysis of GPS travel data.* University of Minnesota.

Huff, J. O., & Hanson, S. (1986). Repetition and Variability in Urban Travel. *Geographical Analysis, 18*(2), 97-114. doi:10.1111/j.1538-4632.1986.tb00085.x

Ishaq, R., Bekhor, S., & Shiftan, Y. (2012). *Flexible model structure for discrete-choice models.* Paper presented at the Proceedings of the Institution of Civil Engineers-Transport.

Koppelman, F. S., & Bhat, C. (2006). *A self instructing course in mode choice modeling: multinomial and nested logit models*: FTA.

La Paix, L., Chowdhury, M. S. A., & Geurs, K. T. (2017). *Panel data for modelling dynamics in duration of outdoor leisure activities.* Paper presented at the International Choice Modelling Conference 2017.

Lepp, A., & Gibson, H. (2008). Sensation seeking and tourism: Tourist role, perception of risk and destination choice. *Tourism Management, 29*(4), 740-750. doi:http://dx.doi.org/10.1016/j.tourman.2007.08.002

Limanond, T., & Niemeier, D. A. (2003). Accessibility and Mode-Destination Choice Decisions: Exploring Travel in Three Neighborhoods in Puget Sound, WA. *Environment and Planning B: Planning and Design, 30*(2), 219-238. doi:10.1068/b12846

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

Mishra, S., Wang, Y., Zhu, X., Moeckel, R., & Mahapatra, S. (2013). *Comparison between gravity and destination choice models for trip distribution in Maryland.* Paper presented at the Transportation Research Board 92nd Annual Meeting.

Mokhtarian, P. L., & Salomon, I. (2001). How derived is the demand for travel? Some conceptual and measurement considerations. *Transportation Research Part A: Policy and Practice, 35*(8), 695-719. doi:http://dx.doi.org/10.1016/S0965-8564(00)00013-6

Montini, L., Antoniou, C., & Axhausen, K. W. (2016). Route and mode choice models using GPS data. *Arbeitsberichte Verkehrs-und Raumplanung, 1204*.

Muller, P. O. (2004). Transportation and urban form-stages in the spatial evolution of the American metropolis. In S. Hanson & G. Giuliano (Eds.), *The geography of urban transportation* (3rd ed., pp. 59-85): New York : Guilford Press.

Nadine Schuessler, I., & Axhausen, K. W. (2009). Accounting for Route Overlap in Urban and Sub-urban Route Choice Decisions Derived from GPS Observations.

Newman, J. P., & Bernardin, V. L. (2010). Hierarchical ordering of nests in a joint mode and destination choice model. *Transportation, 37*(4), 677-688. doi:10.1007/s11116-010-9277-x

Outwater, M. L., Bradley, M., Ferdous, N., Bhat, C., Pendyala, R., Hess, S., . . . LaMondia, J. (2015). *Tour-Based National Model System to Forecast Long-Distance Passenger Travel in the United States.* Paper presented at the Transportation Research Board 94th Annual Meeting.

Pas, E. I., & Sundar, S. (1995). Intrapersonal variability in daily urban travel behavior: some additional evidence. *Transportation, 22*(2), 135-150.

Pendyala, R. (1998). Causal analysis in travel behaviour research: a cautionary note. *Travel behaviour research: updating the state of play*, 35-48.

Pendyala, R., & Pas, E. (2000). *Multi-day and multi-period data for travel demand analysis and modeling* (0097-8515). Retrieved from http://onlinepubs.trb.org/onlinepubs/circulars/ec008/workshop_b.pdf

Pozsgay, M., & Bhat, C. (2001). Destination choice modeling for home-based recreational trips: analysis and implications for land use, transportation, and air quality planning. *Transportation Research Record: Journal of the Transportation Research Board*(1777), 47-54.

Scarpa, R., & Thiene, M. (2005). Destination choice models for rock climbing in the Northeastern Alps: a latent-class approach based on intensity of preferences. *Land economics, 81*(3), 426-444.

Schlich, R., & Axhausen, K. W. (2003). Habitual travel behaviour: Evidence from a six-week travel diary. *Transportation, 30*(1), 13-36. doi:10.1023/a:1021230507071

Schlich, R., Schönfelder, S., Hanson, S., & Axhausen, K. W. (2004). Structures of leisure travel: temporal and spatial variability. *Transport Reviews, 24*(2), 219-237.

Schönfelder, S., & Axhausen, K. W. (2003). Activity spaces: measures of social exclusion? *Transport policy, 10*(4), 273-286.

Schönfelder, S., & Axhausen, K. W. (2010). *Urban rhythms and travel behaviour: spatial and temporal phenomena of daily travel*: Ashgate Publishing, Ltd.

Schüssler, N., & Axhausen, K. W. (2009). *Accounting for similarities in destination choice modelling: A concept*: Citeseer.

Simma, A., Schlich, R., & Axhausen, K. W. (2002). *Destination choice modelling for different leisure activities*: IVT, ETH Zürich.

Simmons, R., Browning, B., Zhang, Y., & Sadekar, V. (2006). *Learning to predict driver route and destination intent.* Paper presented at the Intelligent Transportation Systems Conference, 2006. ITSC'06. IEEE.

Stopher, P., & Zhang, Y. (2010). *Is travel behaviour repetitive from day to day?* Paper presented at the Australasian Transport Research Forum (ATRF), 33rd, 2010, Canberra, ACT, Australia.

Tarigan, A., Fujii, S., & Kitamura, R. (2012). Intrapersonal variability in leisure activity-travel patterns: the case of one-worker and two-worker households. *Transportation Letters, 4*(1), 1-13. doi:10.3328/TL.2012.04.01.1-13

Thill, J.-C. (1992). Choice set formation for destination choice modelling. *Progress in Human Geography, 16*(3), 361-382.

Thomas, T., & Geurs, K. (2016). *Mode choice dynamics: an exploration of intrapersonal mode choice variation using the Dutch Mobile Mobility Panel.* Paper presented at the Mobile Tartu 2016, Estonian Biocentre (Omicum); Riia 23b-105, Tartu.

Train, K. E. (2009). *Discrete choice methods with simulation*: Cambridge university press.

Transit Modeling Update. Trip Distribution Review and Recommended Model Development Guidance Retrieved from http://www.fsutmsonline.net/images/uploads/reports/Trip_Distribution_Review_and_Recommended_Improvements_.pdf

Wang, Y., Wu, B., Dong, Z., & Ye, X. (2016). A Joint Modeling Analysis of Passengers' Intercity Travel Destination and Mode Choices in Yangtze River Delta Megaregion of China. *Mathematical Problems in Engineering, 2016*.

Wu, L., Zhang, J., & Fujiwara, A. (2011). Representing tourists' heterogeneous choices of destination and travel party with an integrated latent class and nested logit model. *Tourism Management, 32*(6), 1407-1413. doi:http://dx.doi.org/10.1016/j.tourman.2011.01.017

Yáñez, M. F., Cherchi, E., Heydecker, B. G., & de Dios Ortúzar, J. (2011). On the treatment of repeated observations in panel data: efficiency of mixed logit parameter estimates. *Networks and Spatial Economics, 11*(3), 393-418.

Ye, X., Cheng, W., & Jia, X. (2012). Synthetic Environment to Evaluate Alternative Trip Distribution Models. *Transportation Research Record: Journal of the Transportation Research Board*(2302), 111-120.
Zimmermann, M., Mai, T., & Frejinger, E. (2017). Bike route choice modeling using GPS data without choice sets of paths. *Transportation Research Part C: Emerging Technologies, 75*, 183-196. doi:http://dx.doi.org/10.1016/j.trc.2016.12.009

Appendices

Appendix A: Data Processing, Cleaning and Enriching

One of the first and most important steps in any data processing task is to verify that data values are correct or, at the very least, conform to some a set of rules (Cody, 2008). Despite of being a unique dataset, DMMP has issues like multiple entries, unidentified destination, wrong records, broken trips, etc. These issues are identified and corrected to avoid duplicated or missing information that can produce incorrect or misleading statistics. In terms of dealing with the error and missing data, and also to enrich the data, repetitiveness of the travel behavior considered. For example, Simmons *et al.* (2006) stated that driving is a routine activity for many people where they drive to the same destinations (e.g. work) using the same routes (along with a small set of routes) on a regular basis. Similarly, people can shop on different days or at different times, however they might choose to shop from the same store(s). For this study, enriching the data is done mainly based on spatial and temporal repetition.

Missing Locations are replicated from the maximum repeated location for a certain trip purpose per individual. However, not all the trips are repetitive. Home trips, work trips can be repetitive predictively but other trips, like walking around, travelling, leisure can vary. Even, for the work trips, one person can have multiple work location. Considering these constraints, only home trips were replicated from the location based repetition.

Unidentified trip purposes are replicated with the same assumption as for the missing locations – only trips towards home were replicated. There were 3.59% trips, where the trip purposes were unidentified and 3.35% trips, where the trip purpose was empty. Among these, only 0.21% of the trips were identified as home trips and are replicated.

Although the **weather data** (weather condition, rain and temperature) were available for the MoveSmarter trip detection, however, that information was missing in the revised follow-up data that is filtered for analysis. To deal with this situation, first, weather information was retrieved from the original MoveSmarter data. Afterwards, still there was a lot of missing values and those were filled by the maximum reported weather situation of that day.

Negative Travel Time is found which could occur for multiple reasons. While, travel time is calculated by subtracting the arrival time from the departure time, sometimes, respondents forgot to correct departure/arrival time, and therefore wrong departure/arrival time is retrieved. Such cases were handled manually by two different ways: First, based on the temporal dimension and trip repetition, it was possible to update the departure/arrival time. And secondly, if it does not match in any pattern and found as an outlier, those trips were removed from the dataset and not further considered for analyses.

Along with negative travel time, there were issues with **negative travel time differences**. This means, when a respondent completed two trips and reported same arrival time, the difference between the departure time of the next trip and arrival time of the previous trip becomes negative. In such cases (approximately 600), the trips were taken care manually. Some of them were found as duplicate and therefore removed. Some of them were corrected based on the respondents' trip repetition, travel pattern.

Zero Travel Time – means the departure and the arrival time is the same were found 126 times precisely. These records were removed and not further considered in the analyses.

Duplicate Entries were removed at the beginning of data mining process. However, sometimes there were same trips that were detected by the MoveSmarter and revised by the respondents with small change in either departure or arrival time but both were present as the revised data by respondents. In such case, revised trips were kept in the dataset for further analyses.

Trip Chaining was an important manipulation of the dataset. Sometimes, there are trips that had a break, perhaps, for a traffic signal or some other uncertainties with same trip purpose. These broken trips were joined based on two assumptions. First, when the trip purpose and mode are same, and the departure time of the next trip is equal to the arrival time of the previous trip, are considered as continuous. The same assumption was replicated for the destination postcodes as well. Secondly, if the trip purpose and mode are same and the difference between the departure time of the next trip and the arrival time of the previous trip is less than or equal to 10 minutes, are also considered as continuous trips. The argument behind this assumption is, sometime respondents take a break for refueling or for taking food or may be congestions or some other uncertain circumstances.

No Trips Respondents sometimes misunderstand the data note strategy. The purpose of the follow-up survey was to make sure if the data collection by MoveSmarter is good, if not, to revise. However, respondents sometimes revised the data by adding a trip mentioning that s/he was in the house. Again, sometimes a walk inside the house is reported by the app and somehow respondents also approved it as a trip. Along with these, sometimes some small trips were recorded at the same location, which is also weird. This sort of records is coded as no trip.

Foreign Trips are coded by the match of the postcodes – the format number of the postcodes are unlike Dutch postcodes. Afterwards, GPS coordinates outside the Netherlands are also coded as foreign trips.

To define the destination characteristics, **land use information** is imported from the Open Street Map information. First, each PC6 area is defined with a certain land use type which covers maximum areas of the corresponding postcode. This procedure is replicated also for the PC5 and PC4 level. The reason for considering more aggregated level is the lack of PC6 data for all the records. Therefore, when a PC6 is not available and PC5 or PC4 is available, still the type of the land usage can be defined.

Lastly, there are only 2% of the data named as unknown since they are missing the destination postcode although the departure postcodes are available. These data are excluded from the analysis.

Appendix B: Data availability and Selection

Segmentation and Trip Purposes

Table A 1: Available trip purpose on each segment

Fixed destinations	Flexible Destinations
To work	Shopping (grocery and others)
To home	Free time/leisure
Education/seminar	Visit
Business trip	Hiking/sightseeing walk
Bringing away or picking up	Sport/hobby
Personal care	Going out
	Others

Available Variables and Sources

Socio-Economic Characteristics

- Gender, Age, Occupational status, Educational status, Marital status, Urbanity level of the residential area, No. of kids in HH, Income, Etc.
- Source: DMMP dataset which has been retrieved from the LISS panel

Trip Characteristics

- Trip purpose, departure and arrival time, travel time, mode usage, etc.
- Source: DMMP dataset

Spatial Variables

- **land-use types**
- allotments, cemetery, commercial, farm, forest, grass, heath, industrial, meadow, military, nature reserve, orchard, park, quarry, recreational ground, residential, retail, scrub, vineyard
- Source: Open Street Map (http://download.geofabrik.de/europe/netherlands.html),
- Description of the variables can be found at: http://wiki.openstreetmap.org/wiki/Key:landuse

- **Built environment**
- distance to doctor, super shop, daily grocery, warehouse, café, café plus food, restaurant, hotel, day-care for kids/babies, place where students can Stay and do their homework after school hours, fire department, onramp of highway, train station, transfer point for public transport, swimming pool, ice skating, library, cinema, sauna, tanning bed, activities (sport, culture etc.).
- Source: Data were collected from CBS from https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische%20data/wijk-en-buurtkaart-2015

- **Job accessibility**
- Bike and ride (BnR), walk and ride (WnR), Car, etc.
- Source: The values are estimated using Lisa dataset 2014 for jobs. This is a census of all companies in the Netherlands. The car accessibility is calculated using historical speed profile data from TomTom and also for the period of 2014. The public transport (WnR) is a GTFS model for which, the data came from 9292 ov. The bike and ride is an expanded model that has bike speed data from the fietselweek data along with GTFS. That is done considering the shortest path between just taking the bike (trips more than 200 m but less than 30 mins - this 30 mins came from analysis from the OVIN data), the regular GTFS model, and using the bike as an access more to train stations. The times were calculated using network analyst and for the Netherlands all of them used a log logistic distance decay function.

Others

- Weather condition (clear air, cloudy, rainy, fog, etc.), Temperature, Rain,
- Weekdays (Workdays, weekend)
- Source: DMMP dataset

Appendix C: Statistics

First, a detailed descriptive analysis is carried out for all the trip char., SE, spatial variables. The descriptive statistics revealed the distribution of the data (trips) over those variables and as well as for each defined destination type alternative. Afterwards, SRI and TRI distribution over the selected trip characteristics, SE is plotted and elaborated the intrapersonal variation more.

Descriptive Statistics

Socio-Economic Characteristics

The distribution of the respondents with different important socio-economic variables are shown below in Table A 2. The variables that are presented in the table, are categorical. The percentages of the alternatives are estimated with each type of destinations.

Table A 2: Descriptive of SE (%)

		% in data	Fixed destinations				Flexible destinations				Fixed destinations				Flexible destinations			
			MV	MuV	EV	VO	MV	MuV	EV	VO	Very high	high	medium	low	Very high	high	medium	low
Gender	Female	51	30	23	19	28	27	21	26	26	24	25	25	26	29	29	25	17
	Male	49	31	21	19	29	28	20	26	26	24	25	26	25	28	30	25	18
Age	15-24	11	31	23	16	30	27	21	26	26	26	23	25	26	28	30	25	16
	25-34	14	30	23	19	28	28	18	28	26	22	25	26	26	26	29	25	20
	35-44	21	31	24	17	29	28	20	26	26	23	25	26	26	27	30	26	18
	45-54	24	32	20	18	29	28	21	25	27	23	25	26	26	29	30	23	18
	55-64	26	31	22	20	28	27	21	26	26	24	25	25	26	29	29	25	17
	>=65	15	35	17	20	28	28	21	26	26	25	25	25	25	33	30	23	13
Occupational status	Employed	54	30	23	18	29	28	20	26	26	23	26	26	26	27	29	26	18
	Self-employed	6	32	20	21	27	30	16	28	26	26	24	25	25	31	32	20	17
	Unpaid work	3	32	19	24	24	28	20	28	25	24	27	24	24	32	29	21	18
	Retired	16	33	19	20	28	27	21	25	26	25	25	24	26	33	30	23	15
	Searching for job	8	32	20	19	28	28	24	22	27	25	23	26	26	31	30	22	17
	Housekeeping	5	32	21	19	29	26	22	26	26	25	24	24	27	31	29	23	17
	In school	10	31	24	16	30	26	22	26	25	25	23	26	26	29	31	26	14
	Incapacitated	7	35	18	20	27	28	19	25	27	26	24	26	25	32	31	21	16
	Others	2	33	17	21	29	27	23	23	27	23	27	27	23	31	27	23	19
Marital status	Married	53	31	21	20	29	27	20	27	26	24	25	25	26	29	29	25	17
	Separated	1	27	27	18	27	27	27	18	27	25	25	25	25	20	30	20	30
	Divorced	14	31	21	20	28	27	22	24	27	24	24	27	25	30	28	23	18
	Widow/Widower	3	30	23	21	26	28	21	26	26	24	26	26	24	27	27	24	22
	Unmarried	31	30	24	17	29	28	21	25	26	24	25	26	26	28	30	25	17
Education	Basic education	6	34	14	21	31	28	19	27	26	26	22	26	27	32	32	22	15
	VMBO	16	33	22	18	27	27	20	27	26	25	25	25	25	32	30	25	13
	HAVO/HBO	16	31	22	17	30	28	21	26	25	26	24	25	26	29	30	26	16
	MBO	25	30	23	18	29	27	21	26	26	23	26	26	25	28	29	24	19
	HBO	27	30	22	21	27	28	20	26	26	24	25	25	26	29	29	24	18
	WO	12	29	23	19	29	27	21	26	26	23	25	26	26	26	28	27	18
Urbanity	Urban	64	30	22	20	28	27	21	26	26	24	25	26	26	28	29	25	18
	Semi Urban	20	30	23	18	28	27	20	27	26	25	25	25	25	29	30	24	16
	Not Urban	16	31	21	18	29	28	20	25	27	25	24	25	26	30	29	23	17

Table A 2 revealed that most of the time the individuals' participation in each alternative of each destination type is not substantial, however, in percentile distributed flexible destination alternatives, participation of the individuals having SRI below 25 percentile has a lower share. The gender distribution in the sample is almost equal: female (51%) and male (49%). Over the age, more people seem to be attracted in alt. 4 of flexible destination and vice versa for the alt. 1. The Same observation can be seen for the retired and employed people. This is relevant since; elderly people (and also retired people) repeat same locations for the activities towards flexible destination whereas younger people tend to explore more. 54% respondents are employed, 16% are retired and 10% are found in the school. Most of the respondents are in married (53%) and unmarried (31%). Only 6% respondents are found have just basic education and lastly, most of them (64%) are found living in urban area.

On the other hand, Table A 3 shows the descriptive statistics of the trip distribution for the same socio-economic variables as Table A 2 and the percentages are per population segment.

Table A 3: Trip distribution of SE vs alternatives

		% in dataset	Fixed destinations				Flexible destinations				Fixed destinations				Flexible destinations			
			MV	MuV	EV	VO	MV	MuV	EV	VO	Very hig	high	medium	low	Very high	high	medium	low
G	Female	51	77	11	3	9	28	16	22	34	25	26	25	24	22	27	23	28
	Male	49	76	10	4	10	28	18	20	34	24	25	27	24	24	24	24	28
Age	15-24	11	76	13	3	9	30	14	24	32	26	22	28	24	23	30	24	23
	25-34	14	74	13	3	10	25	12	27	35	19	27	28	26	16	24	25	35
	35-44	21	76	10	3	11	28	16	24	32	24	25	25	26	19	25	24	33
	45-54	24	76	11	3	9	30	19	18	33	19	26	29	26	24	24	22	30
	55-64	26	78	9	4	9	27	19	19	35	28	25	25	22	25	26	26	23
	>=65	15	80	7	5	9	28	18	21	33	39	25	16	20	28	28	22	22
Occupational status	Employed	54	76	11	3	10	27	15	24	34	19	26	29	25	18	24	26	32
	Self-employed	6	71	12	7	10	29	12	25	34	28	19	23	29	26	26	18	30
	Unpaid work	3	85	5	4	6	33	18	15	33	40	23	20	17	33	28	15	24
	Retired	16	78	8	4	9	27	19	19	35	33	27	19	21	28	27	23	22
	Searching for job	8	76	10	5	9	27	19	19	34	32	25	20	23	26	29	24	22
	Housekeeping	5	81	7	4	8	28	24	14	34	41	25	13	21	25	26	23	26
	In school	10	76	13	3	8	27	14	26	32	28	22	27	23	22	30	25	22
	Incapacitated	7	82	6	4	8	34	23	13	30	40	23	18	20	35	27	15	24
	Others	2	76	4	6	13	28	20	20	32	24	30	18	27	30	25	25	21
Marital	Married	53	77	10	4	10	27	17	22	34	27	24	24	25	23	25	24	28
	Separated	1	71	15	2	11	23	30	15	32	16	30	24	30	16	20	22	41
	Divorced	14	77	9	4	10	28	18	20	35	21	28	25	25	24	28	21	27
	Widow/Widower	3	71	17	4	9	31	17	11	41	26	22	26	26	22	20	21	37
	Unmarried	31	77	11	3	9	30	16	23	32	23	26	29	23	23	28	25	25
Education	Basic education	6	81	5	6	9	26	19	21	33	37	26	18	19	27	31	20	21
	VMBO	16	79	9	3	9	29	17	20	34	33	25	21	21	27	28	21	24
	HAVO/HBO	16	75	13	3	9	30	17	20	32	23	26	26	25	23	27	23	27
	MBO	25	77	9	3	10	29	18	19	33	23	27	26	23	24	24	23	30
	HBO	27	75	12	4	10	26	16	24	34	22	24	28	26	21	25	26	28
	WO	12	77	9	3	10	27	15	24	34	21	25	29	26	17	26	26	31
Urba	Urban	64	77	10	4	9	28	18	21	34	23	26	27	24	22	25	25	28
	Semi Urban	20	79	9	3	9	30	16	23	32	30	25	24	22	24	26	24	26
	Not Urban	16	74	10	3	13	29	14	22	35	23	24	24	28	23	27	20	31

It can be seen that, most visited locations dominate the trip percentages for the fixed destination segment. Presumably, these destinations are fixed by purposes and have high share, however different picture can be seen in flexible destination segment. The trip percentages have fluctuation and most of the time visited once alternative is dominating, followed by most visited locations. In the SRI distributed alternatives, quite uniform trip distribution can be observed except few. For instance, mid-aged people seem to have a high share of trips in the lowest repetition index alternative for flexible destinations – clearly a proof of variation seeking travel behavior. This behavior seems to be present also in employed people – 32% of the trip share is in the lowest repetition index category while only 18% in the highest alternative. The scene is just the opposite for the retired people. Along with the respondents' percentages, the trip share represents the same picture for the age and employment status as discussed previous paragraph. The trip share in flexible destination (SRI distributed) alternatives, it can be seen that the percentages of trips in low SRI increases with the educational level, means the variation seeking tendency become stronger with educational level.

Trip Characteristics

It was clear that car is the dominating mode followed by bicycle and walking. Figure A 1 is a more detailed version and at more disagrregated level of the earlier statistics. Thus, it can be seen that still for each destination type and for each alternative, car is dominationg mode followed by bicycle and walking. Train and BTM (Bus-Tram-Metro) have been found a higher trip share for fixed destinations. For example, for multiple visited locations, train is competing with walking and bicycle. Moreover, train and BTM have a notable share in MuV, VO, medium and low SRI defined alternatives for fixed destination.

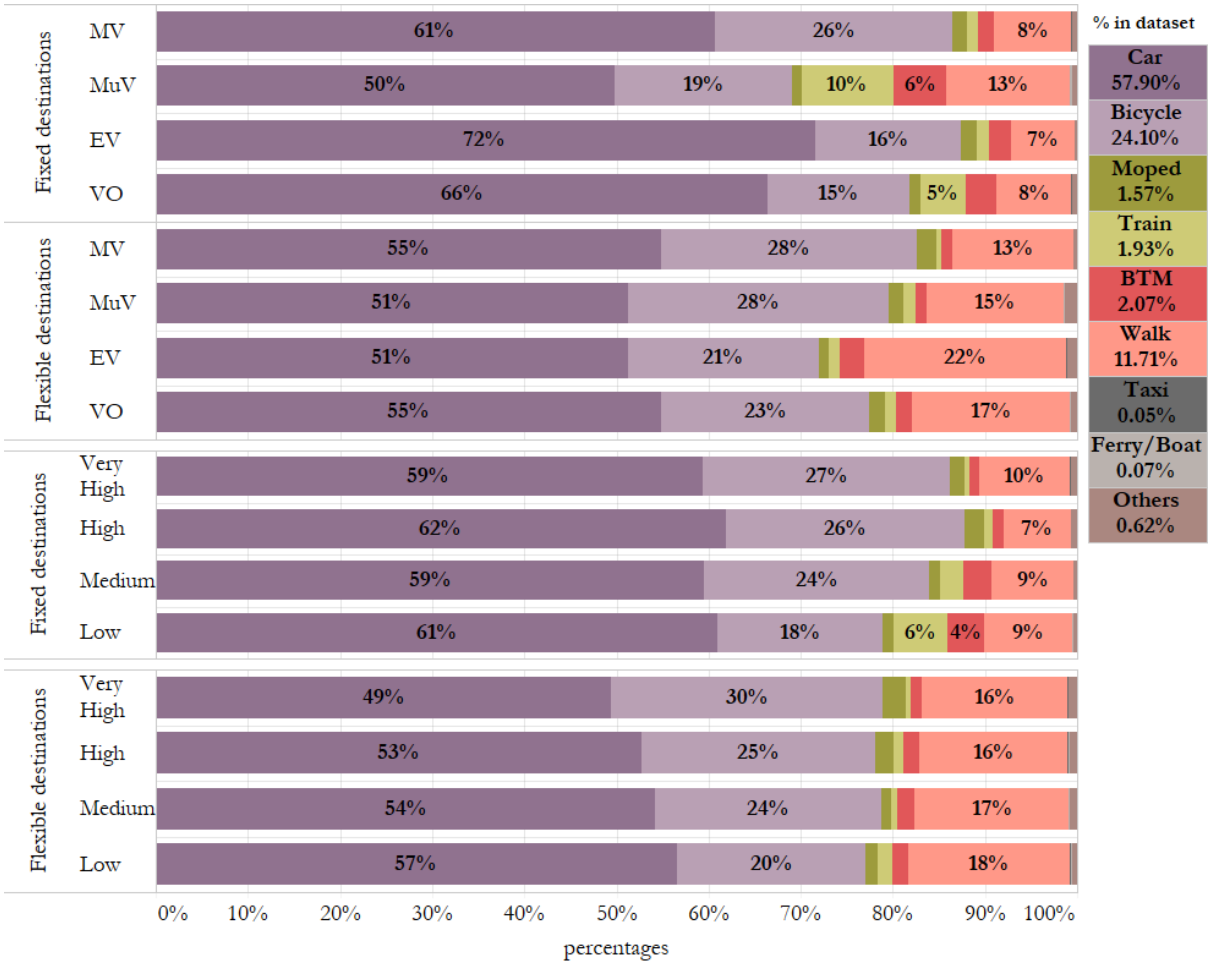

Figure A 1: Mode Choice vs Alternatives

Looking into the departure time of the trips (Figure A 2), the journey towards fixed destinations start from the 5 AM with a peak at 8 AM. On the other hand, there is another peak at 5 PM. These are habitual since the activity types in fixed destinations are mostly work, education, appointment and home trips and usually these trips generate around 8 AM and 5 PM. Nonetheless, trips toward flexible destinations are different. Figure A 2 shows that these trips start usually at 7 AM and quite uniform throughout the day till evening (7 PM), although three small peaks at 10 AM, 1 PM and 7 PM can be seen. This explains, for example, shopping trips usually don't start to early, for a day out departure may occur in the early morning but for recreational activities, trips start usually generates later morning.

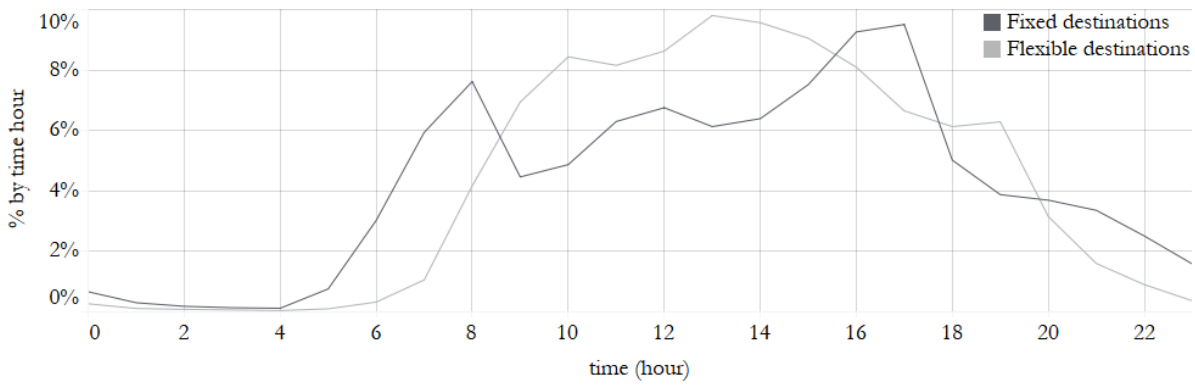

Figure A 2: Departure time

Figure A 3 illustrates the travel time of the respondents for alternative specific destination types. For all the trips towards fixed destination, early and late morning is preferred, more specifically, most- and multiple visited locations are in the early morning and equally- and only visited are in the late morning. On the other hand, flexible destinations have the least share in the early morning, while dominant in the late morning, late afternoon and evening. Surprisingly, medium and low repeated (SRI distributed) alternatives have a high share in the early morning for fixed destinations, which means, although people go out for work or study or appointment in the early morning, however the location can vary. In the flexible destination trips, most departure occurs in the late afternoon and quite uniform throughout the day till evening for the same alternatives.

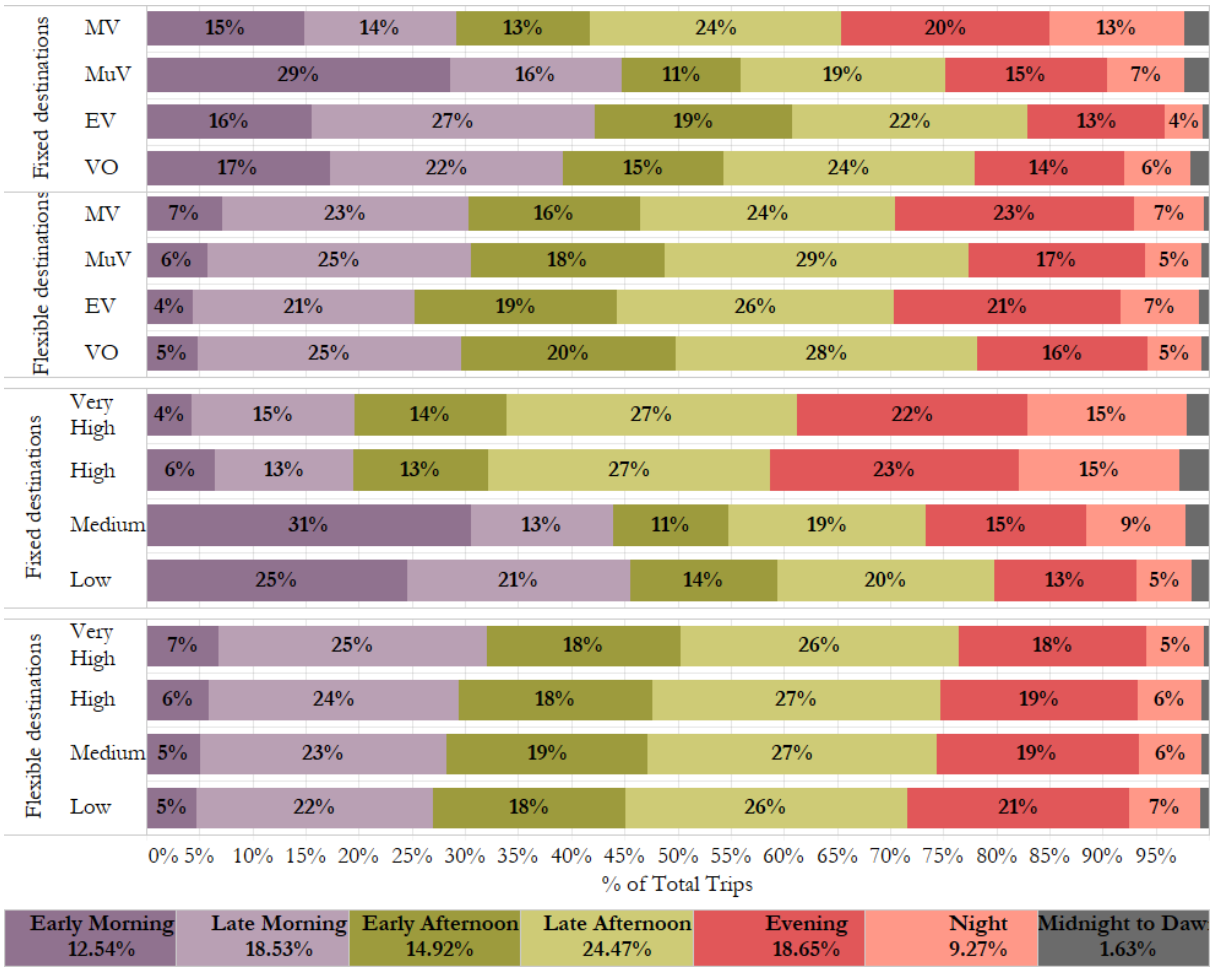

Figure A 3: Departure time vs Alternatives

Looking into the travel time classes, it is found that shorter travel time is dominating each alternative (Figure A 4) – most of the trips (~80%) for each alternative are within half an hour. And the percentages of the trip share are reducing over the increasing in travel time. Longer travel times (1 hour) are performed when the alternatives are less visited (EV, VO) or having lower SRI and particularly for the flexible destination types.

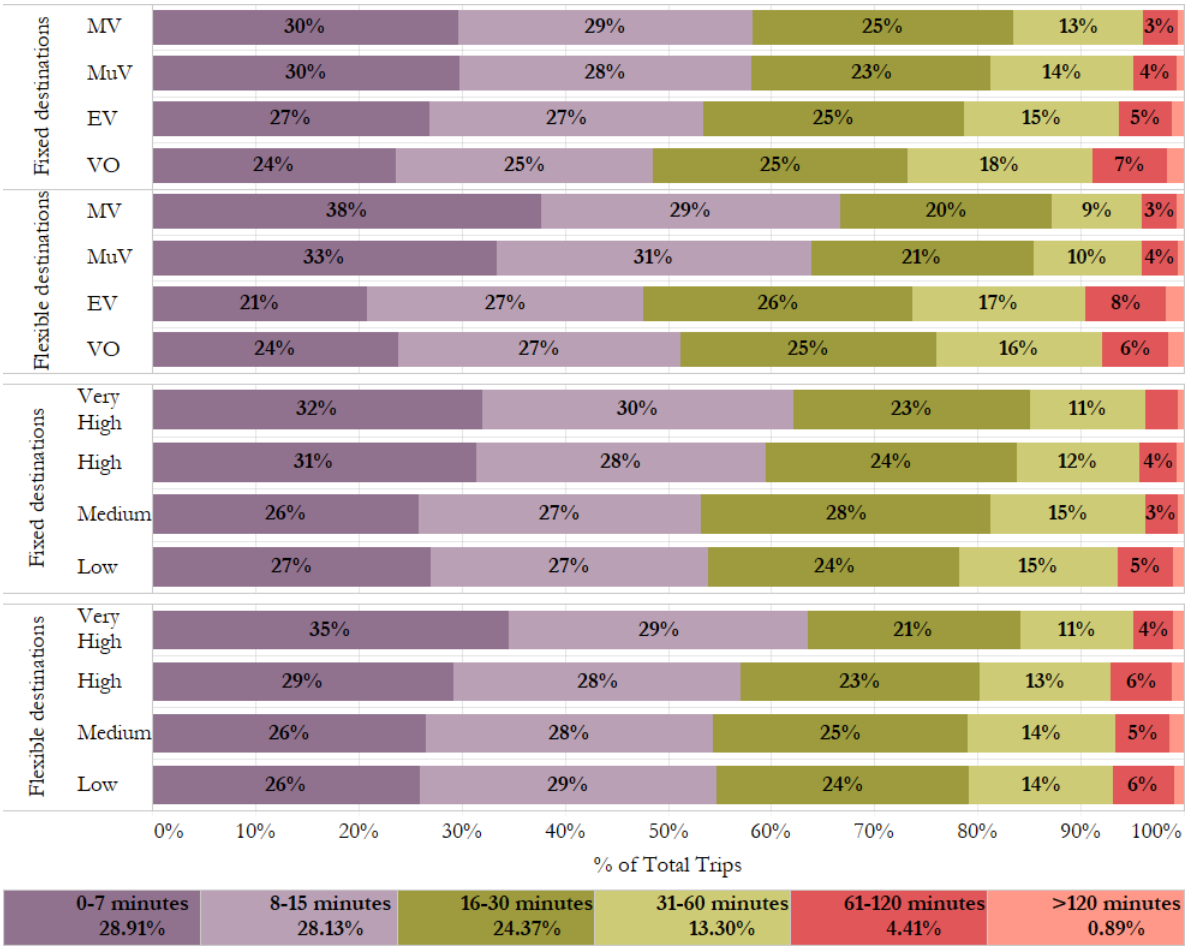

Figure A 4: Travel time classes vs Alternatives

Spatial Variables

As mentioned in Appendix B: Data availability and Selection, a set of land use variables is considered in this study which is retrieved from the open street map data. Table A 4 shows the trip distribution of the land use types over each alternative in each destination type. Some variables are aggregated based on the statistics (see Appendix C: Statistics Dimension Reduction).

Table A 4: Land Use Types vs Alternatives (%)

		allotments	cemetery	commercial & industrial	park, forest & scrub	farm, grass & orchard	meadow & vineyard	military	heath & nature reserve	recreation ground	residential	retail
Fixed	MV	6.19	6.84	16.30	13.08	15.51	7.49	6.56	14.41	0.00	7.78	5.82
	MuV	7.02	6.28	15.41	18.10	11.22	7.06	6.54	7.63	11.58	5.87	3.29
	EV	12.55	5.43	21.37	14.46	11.92	8.16	6.82	7.76	0.00	7.43	4.11
	VO	5.74	3.98	15.42	14.42	12.35	5.85	9.48	10.60	13.86	6.05	2.27
Flexible	MV	7.86	6.23	5.84	37.45	7.85	4.58	6.93	10.61	0.00	3.85	8.80
	MuV	5.35	6.21	9.08	19.51	14.01	5.30	5.91	18.22	0.00	6.69	9.72
	EV	4.73	6.45	10.35	19.79	12.64	5.79	6.11	13.65	7.76	5.97	6.77
	VO	6.51	7.41	11.01	20.97	12.58	6.46	4.60	13.47	2.36	6.35	8.28
Fixed	Very High	8.62	15.58	4.48	12.27	24.01	7.53	0.00	16.11	0.00	10.67	0.74
	High	7.82	5.12	12.11	11.99	17.76	7.07	0.00	21.55	0.00	12.52	4.07
	Medium	9.28	3.65	25.48	14.70	11.75	4.96	8.28	11.30	0.00	6.65	3.95
	Low	3.63	3.26	16.08	13.22	8.99	6.26	10.45	6.32	23.98	4.04	3.77
Flexible	Very High	11.18	8.37	7.02	30.13	8.84	5.42	6.02	11.68	0.00	4.81	6.53
	High	4.38	6.67	9.28	22.54	11.51	5.27	6.13	14.11	4.47	5.31	10.33
	Medium	3.42	7.44	9.70	26.01	11.09	6.49	4.74	14.72	0.00	4.91	11.46
	Low	4.27	4.46	8.02	24.59	11.24	7.35	7.73	12.71	3.75	4.68	11.20

The table shows that, park, forest and scrub area has a high percentage of share in flexible destinations than fixed. On the other hand, commercial and industrial, and farm, grass and orchard area have a higher share in high repeated fixed destinations. Recreational ground has been found with a significant trip share towards fixed destinations. Repeated fixed destinations have a high trip share in residential areas and flexible destinations have in retail areas.

Three (job) accessibility indicators are considered in this study: car, bike and ride (BnR), walk and ride (WnR). It can be seen from the Figure A 5, for multiple visited locations (MuV) and high repeated locations (Very high, high and medium), getting to the fixed destinations is much unlikely than flexible destinations by relying on the public transport (WnR), followed by bicycle (BnR). This clearly represents that respondents prefer public transport and bicycle accessible area more than car accessible area to reach the flexible destinations.

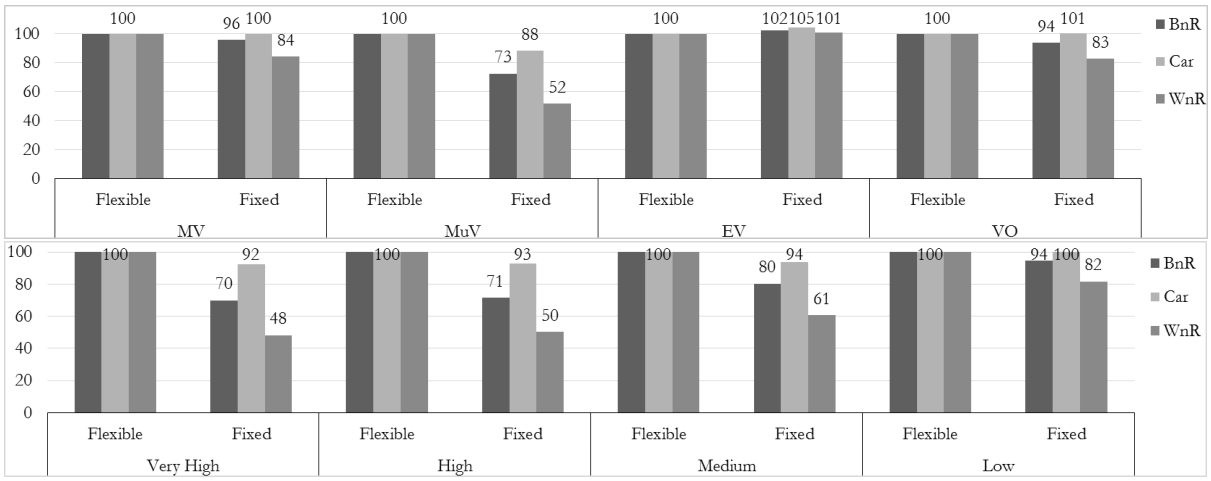

Figure A 5: Accessibility vs Alternatives

Others

Figure A 6 shows that MV destinations are the dominating trips for fixed destination while VO has the highest trip share in flexible destinations. On the other hand, in SRI distributed alternatives, it can be seen that lower repetition index has high trip share during workdays towards fixed destinations. And, in flexible destination, lower repetition index has high share during the weekend.

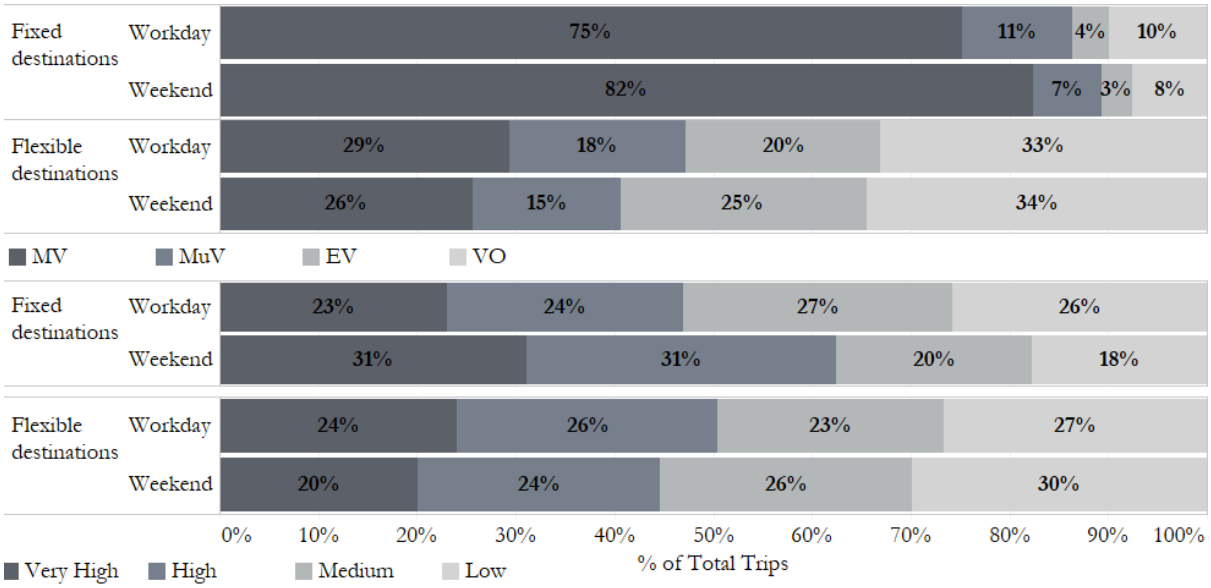

Figure A 6: Weekdays vs Alternatives

Weather seems to have no significant effect on trip distribution towards the alternatives of different destination types (Figure A 7). However, it can be seen that VO and the low SRI in flexible destination have 1% more trips during rain. Although trips towards fixed destinations are expected to have a higher trip share during rain but it seems this is not the case. The reason might be the activity is defining the trip towards a certain destination. For instance, if someone decided to go for outing or shopping, then they carried out the trip even though it's raining.

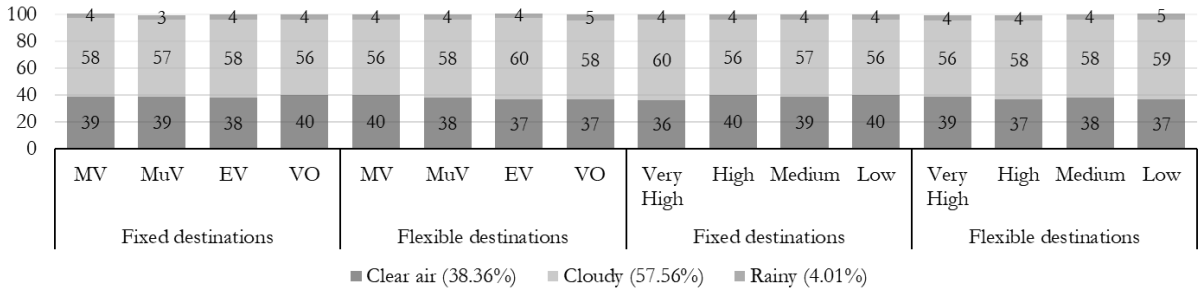

Figure A 7: Weather Condition vs Alternatives

Intrapersonal Variation: SRI and TRI Distribution

The median of the SRI and TRI is already discussed in the Descriptive Statistics section. However, this section describes the distribution of the repetition index (both spatial and temporal) over all available categorical variables (both socio-economic characteristics and trip characteristics).

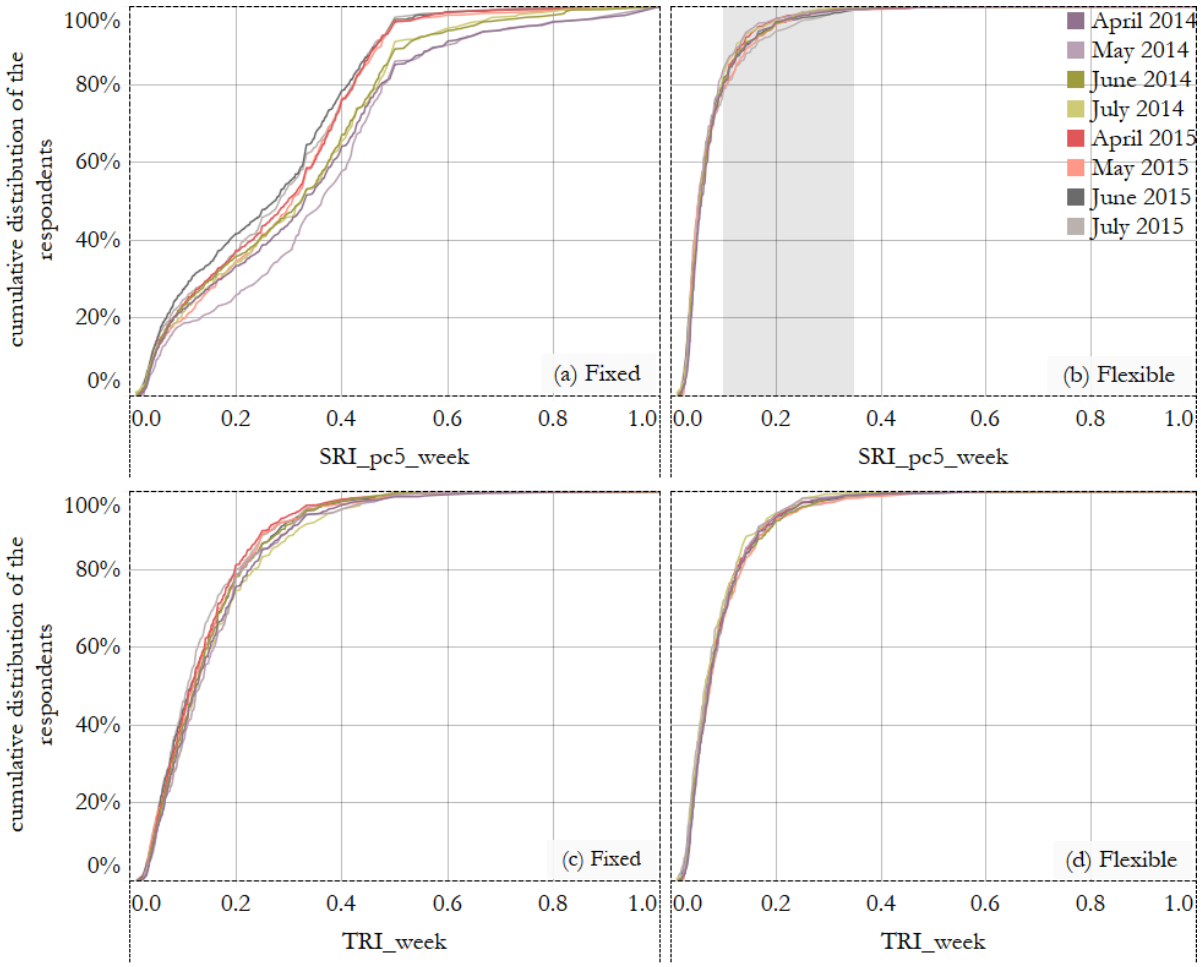

Figure A 8: SRI and TRI distribution over data collection period

First, the indexes are distributed cumulatively over the data collection period. Figure A 8a represents that the repetition index is high in 2014 (particularly in May) than 2015, clearly represents that intrapersonal variation in destination choice is high in 2015 for fixed destination. on the other hand, in flexible destination (Figure A 8b) there is hardly any difference between the SRI distribution, however 0.1-0.35, small differences can be noticed. Interesting is, this is just the other way around than fixed destination. For example, in the mentioned range, the repetition index is higher in 2015 than 2014, which represents that the intrapersonal variation is higher in 2014 for flexible destination than 2015. In temporal repetition index, there are no remarkable differences noticed, although very small variation for the fixed destination trips (Figure A 8).

In Figure A 9, it can be seen that the SRI is much higher than TRI in fixed destination trips, which means the destinations don't vary much while the departure time varies a lot – high intrapersonal variation in departure time in destination choice. On the other hand, in flexible destination trips, this result is vice versa. Thus, the intrapersonal variation in destination choice is higher than the intrapersonal variation in departure time towards flexible destinations.

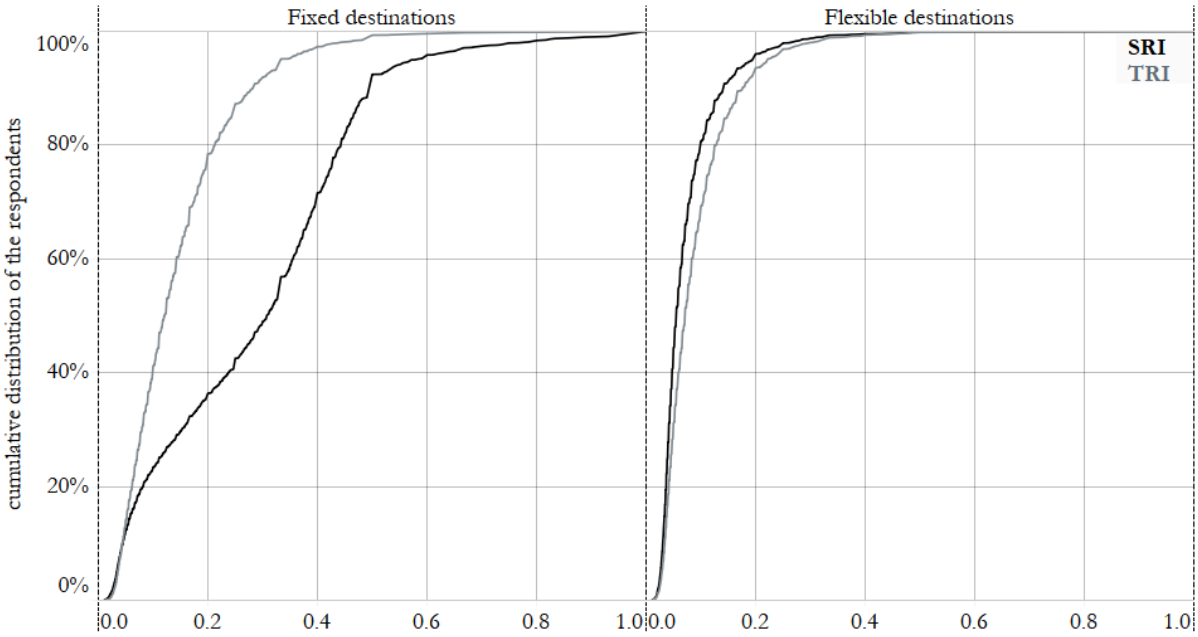

Figure A 9: SRI and TRI differences on destination types

Socio-Economic Characteristics

This section analyses the distribution of the SRI and TRI over crucial socio-economic variables. Figure A 10 shows the repetition indexes distribution for male and females. For fixed destination, the index seems to be higher for female than male respondents. On the other hand, for flexible destination SRI and the both destination type TRI shows that, until a certain value (flexible-SRI ~0.1, fixed-TRI ~0.07, flexible-TRI ~0.12), they aren't different but after that male has high repetition index than female. This means after those values, intrapersonal variation in both destination and departure time is higher for female respondents than males.

The age distribution over destination type and alternatives are already discussed earlier. Figure A 11 revealed the intrapersonal variation in destination choice and departure time over age categories. Notable differences can be observed in fixed destination SRI and flexible destination TRI. For example, elderly people (>55) seems to have higher repetition index in both destination types. Which means that, they have the lowest intrapersonal variation in destination choice. This is habitual since with age the mobility pattern of the people changes and also, elderly peoples are

usually retired from the job. As a result, they show less variation seeking behavior, rather they prefer same familiar destination for the same activity purpose such as, same supermarket, same recreational ground, or same place for treatment, etc. On the other hand, with lower repetition index, young and mid-aged people showed a high intrapersonal variation in the destination choice, although the value overlaps sometimes among the relative groups. In TRI, no remarkable variation is observed in fixed destination segments. However, in flexible destination, TRI has variation and almost similar to the SRI distribution. Consequently, the intrapersonal variation in departure time is found high for the young aged people and low for the elderly people.

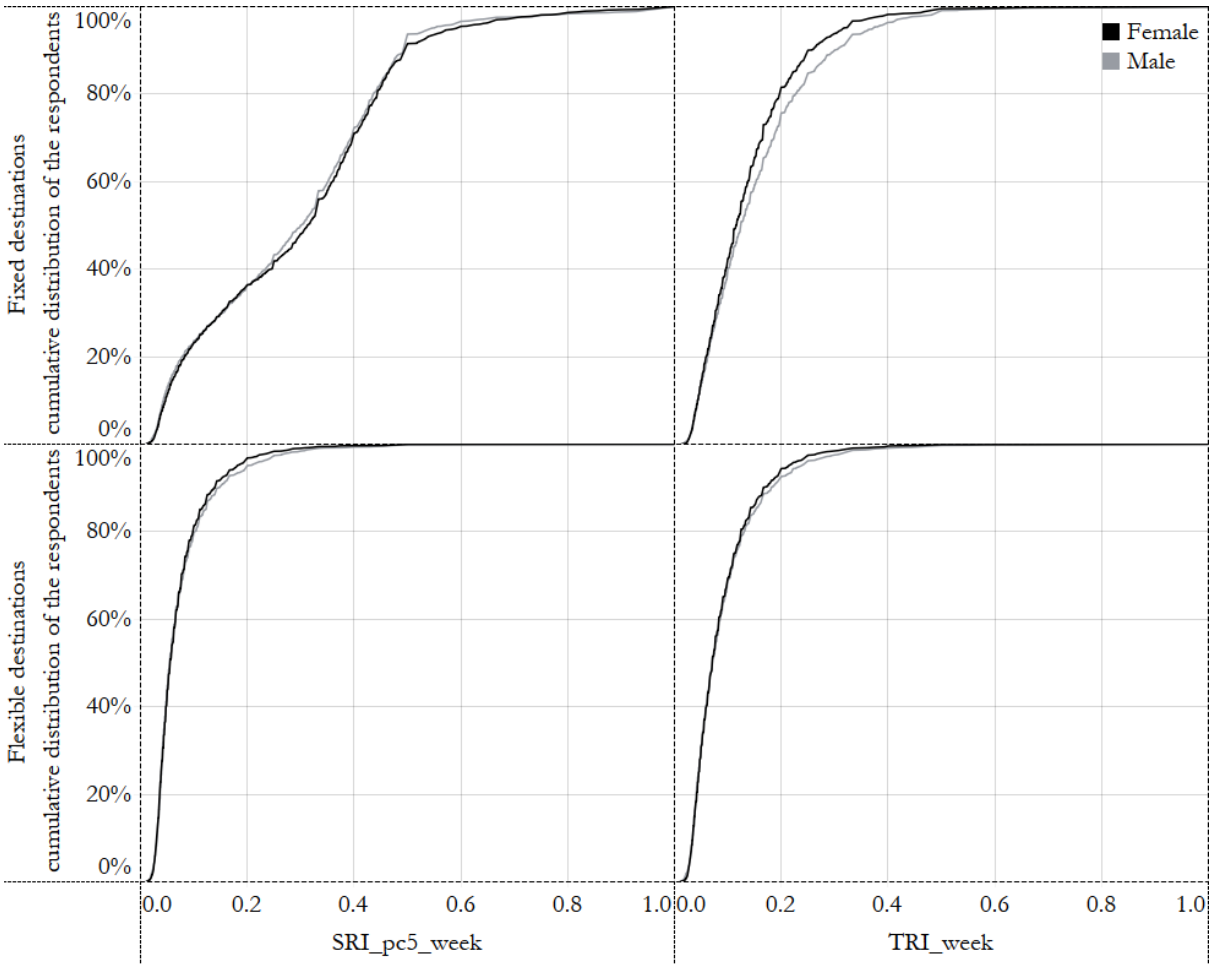

Figure A 10: SRI and TRI distribution over gender

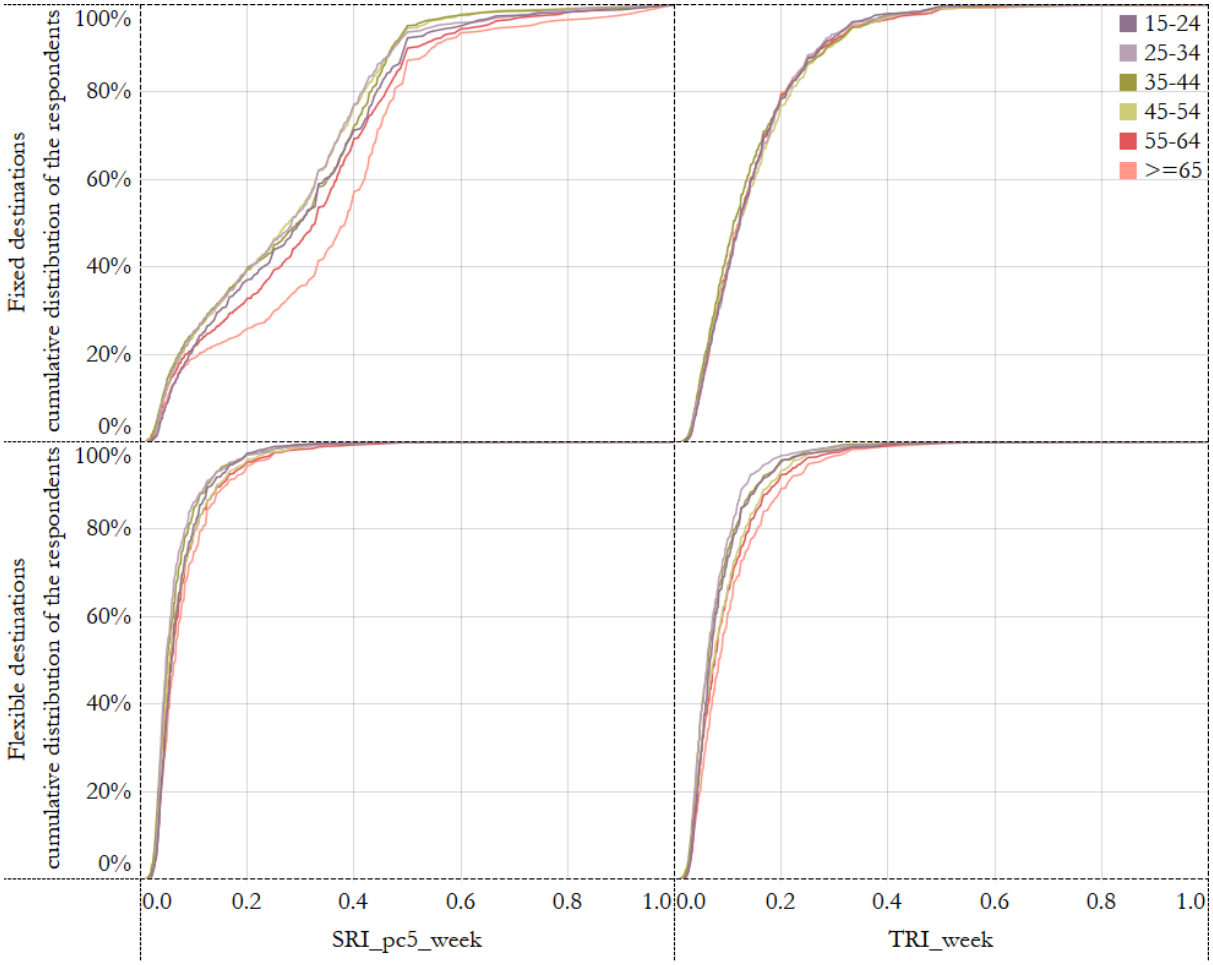

Figure A 11: SRI and TRI distribution over age classes

Although several groups are introduced in the dataset with employment status, however, they are grouped into three categories for easy visualization of SRI and TRI distribution, and analysis: employed (60%), retired (16%) and others (24%). Others include the housekeeping, in school, unpaid work, incapacitated, searching for a job, etc. Figure A 12 reveals that intrapersonal variation in destination choice is high for the employed group for both type of the destinations and lowest for the retired peoples. In fixed-TRI, retired people have the high variation in departure time than employed people which is habitual, since employed people are more time sensitive. On the other hand, flexible-TRI shows the opposite result, means high variation for the employed group and lowest for the retired people. May be retired people repeat the same destination and also at the same time (because of the routine walk, sports or hobby), on the other hand employed people are more relaxed in flexible destinations while they are found time sincere towards fixed destinations.

Figure A 13 describes the SRI and TRI distribution over the educational level. Respondents with basic education and VMBO have high repetition index than other education level, while university degree (WO) has the lowest index value for the fixed destination. Thus, intrapersonal variation in destination choice is high for the university degree people. The Same observation can be observed for flexible-SRI as well, however the differences are very small. At a certain point (~0.05), the difference is even negligible. In TRI distribution, fixed destination picture is very complex to explain, because the difference is too low and they are overlapping a lot, although VMBO has high index most of the time. On the other hand, flexible-TRI shows the similar picture as flexible-SRI, therefore university degree people have high intrapersonal variation in their departure time and lowest for the basic education.

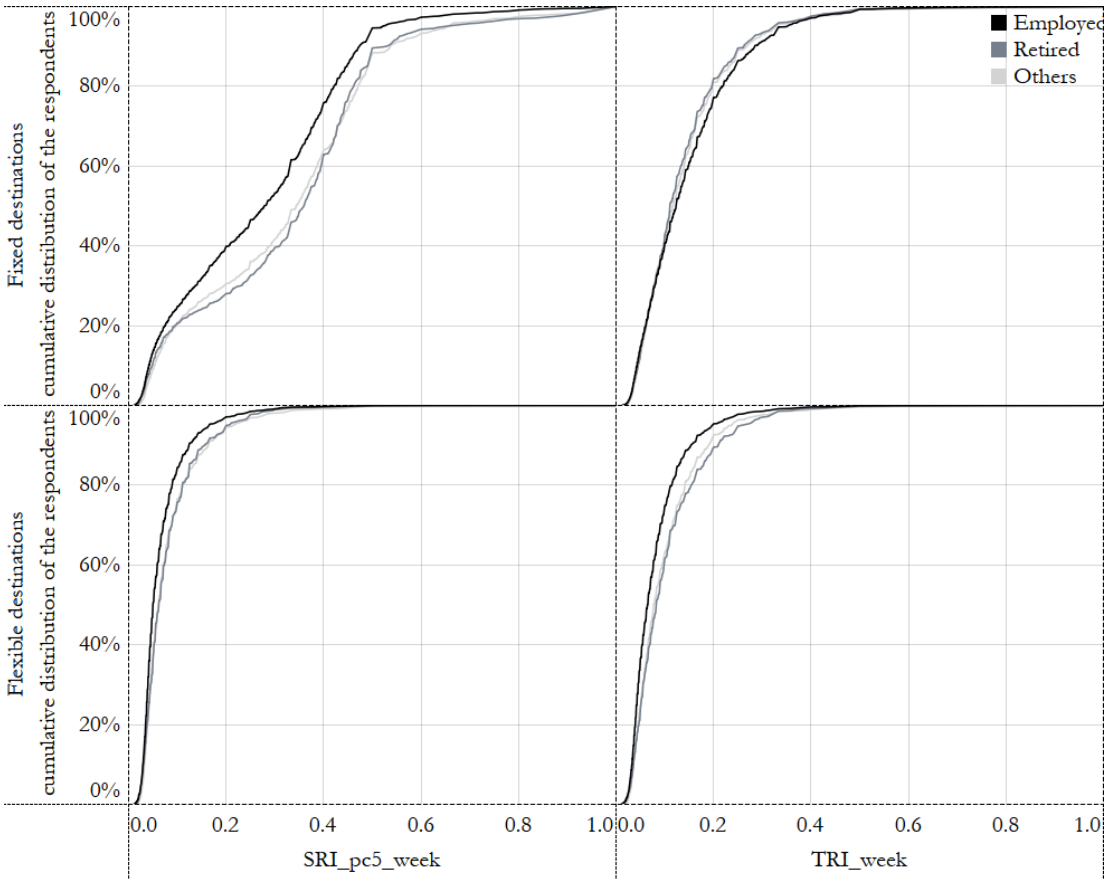

Figure A 12: SRI and TRI distribution over occupational status

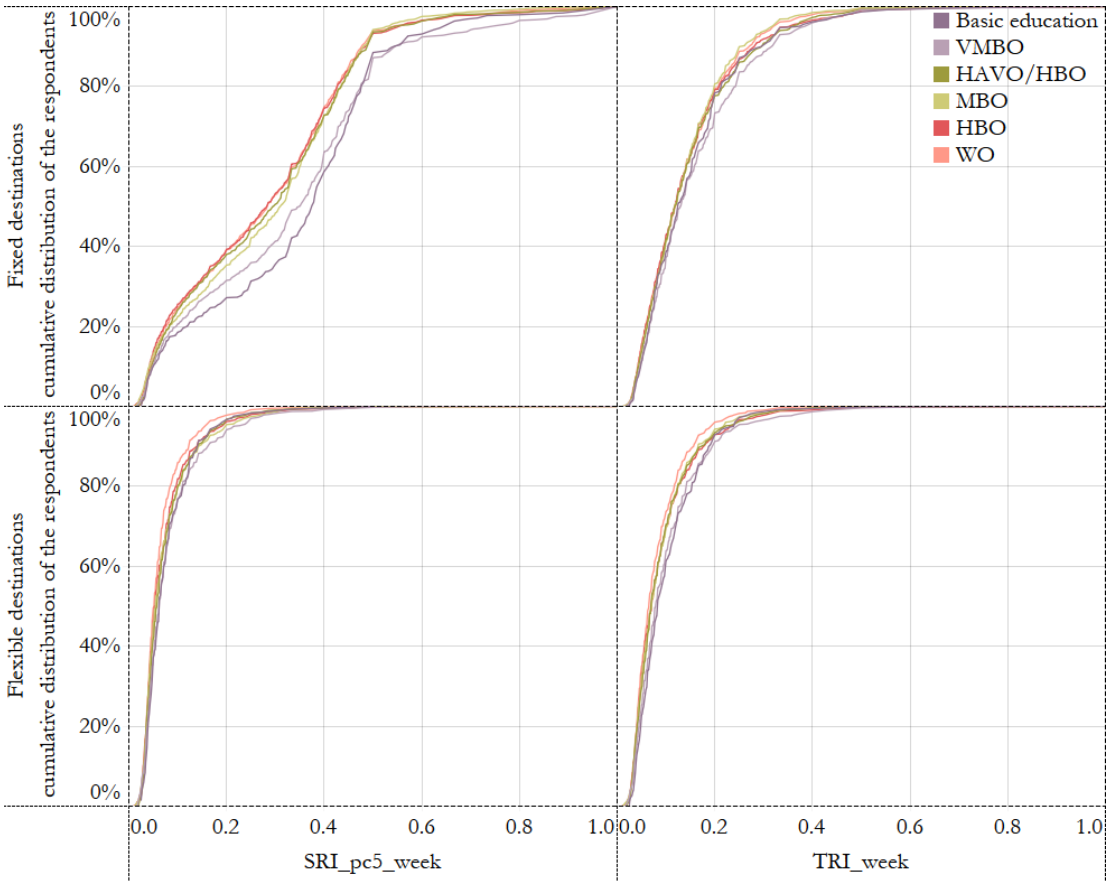

Figure A 13: SRI and TRI distribution over education level

For better visualization, household composition is also grouped into three categories as employment status. Therefore, married (53%), unmarried (31%) and others (16%) groups are defined where others consist of divorced (14%), separated (1%) and widow/widower (3%). Figure A 14 shows that SRI-fixed destination is higher for unmarried respondents until ~0.2. then in between 0.2~0.5, the SRI-fixed is high for married respondents. On the other hand, the SRI-flexible destination is quite uniform and difficult to differentiate. In TRI distribution, fixed destination TRI shows that unmarried respondents have a higher index than married, thus, married people have more variation in starting a journey than unmarried respondents. Flexible destination TRI shows the same picture as SRI of the same segment – not differentiable. However, 0.1~0.2, unmarried respondents have a bit higher variation in departure time than married respondents.

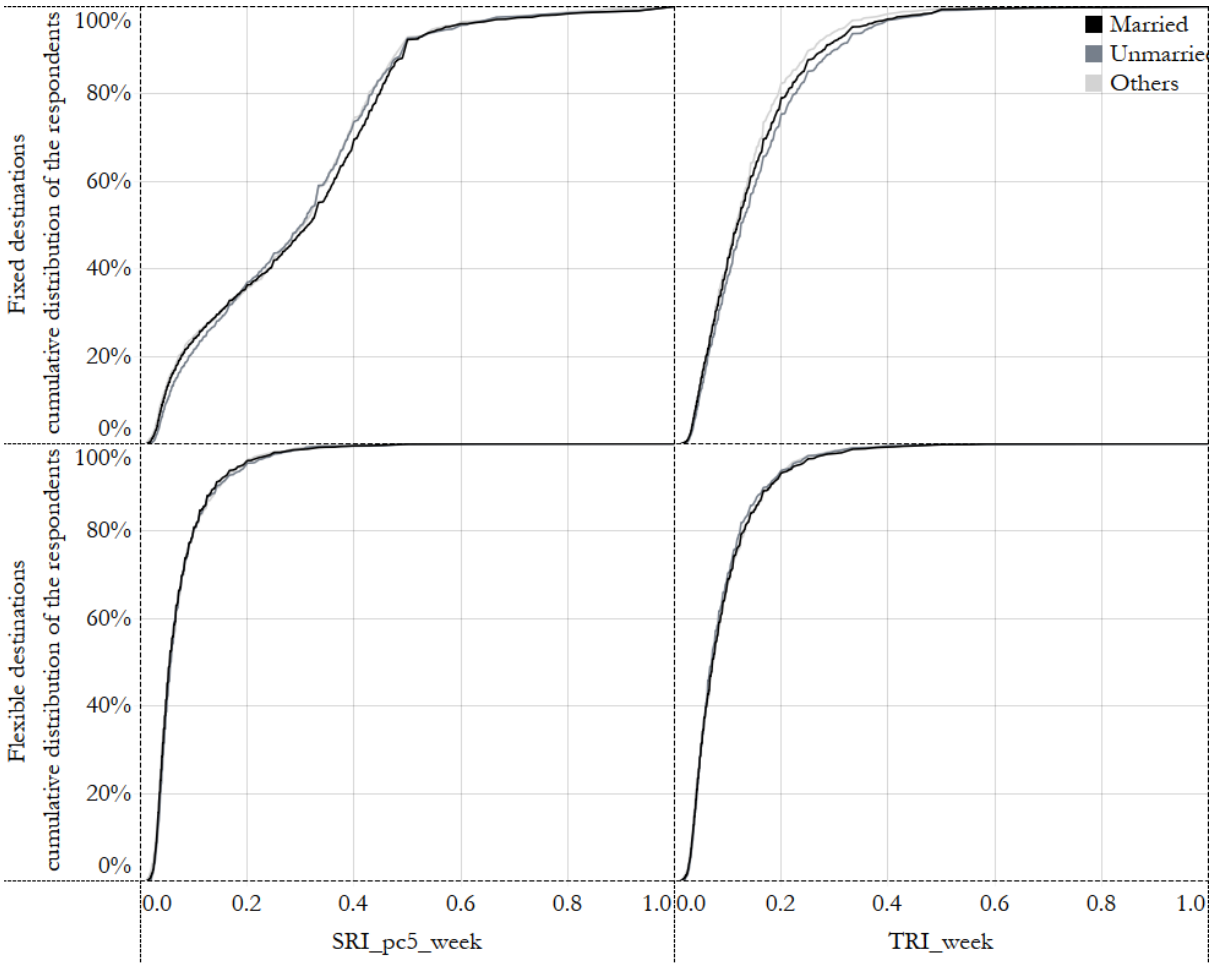

Figure A 14: SRI and TRI distribution over marital status

The distribution of the SRI and TRI over the urbanity level of the residence is shown in Figure A 15. Strongly urban (address density over 2500 km^2), very urban (address density between 1500-2500 km^2) and urban (address density 1000-1500 km^2) are considered as the urban area. Further, two more urbanity indicators are semi urban (address density 500-1000 km^2) and not urban at all (address density <500 km^2). Figure A 15 shows that semi urban areas are highly repetitive followed by urban and not urban area in fixed destinations. Therefore, not urban areas are showing high intrapersonal variation. In flexible destination-SRI, the difference is very hard to notice. Surprisingly the TRI distribution is also alike as SRI of the fixed destination. Nonetheless, TRI distribution of the flexible destinations is different. Departure time from an urban area seems to have higher repetition index than semi urban and not urban. So, people living in urban areas are more time sensitive when they travel to flexible destinations like shop, leisure, sports or going out.

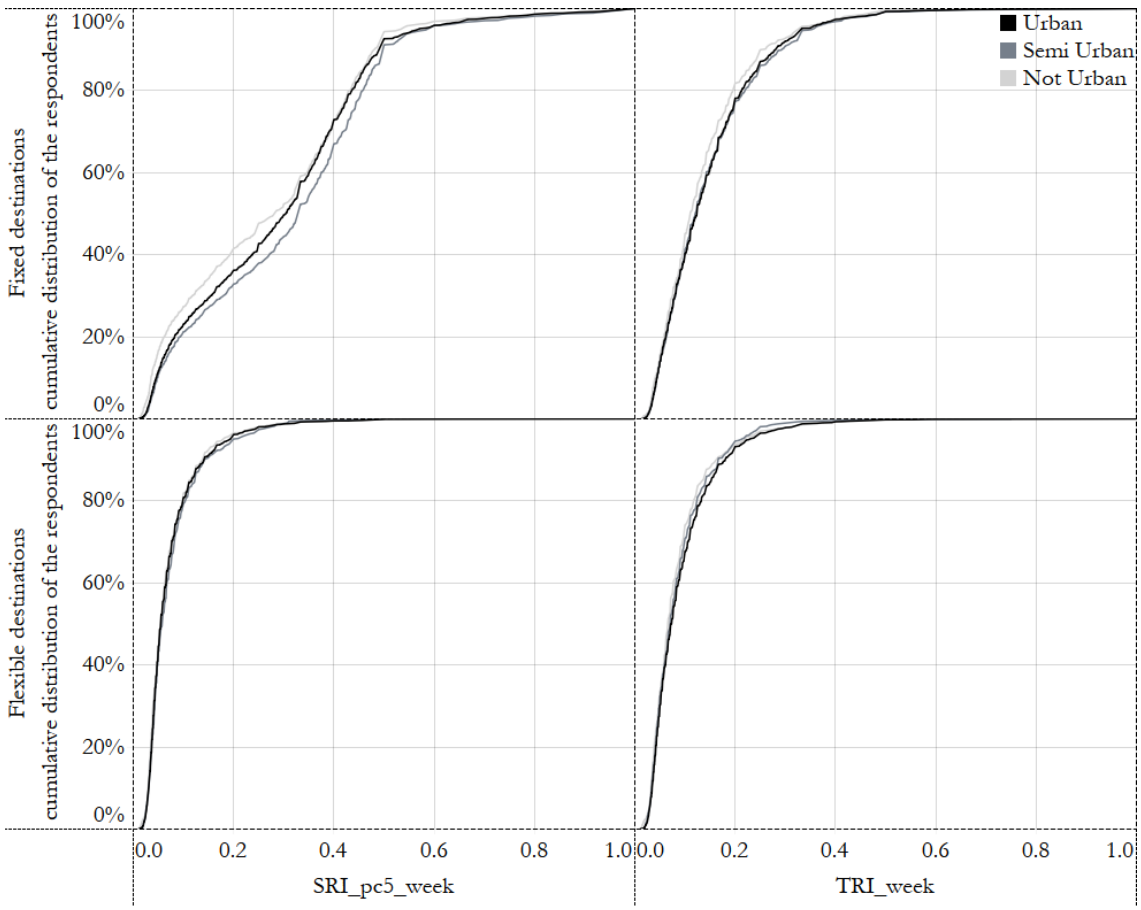

Figure A 15: SRI and TRI distribution over urbanity level of the residence area

Trip Characteristics

Figure A 16 illustrates the distribution of SRI and TRI over mode choice. The available mode usage is already enlisted in Table 4. As expected, bicycle, car is the dominating mode for both fixed and flexible destinations and train, BTM has lower repetition index, which means intrapersonal variation in destination choice is high for public transport than car or bicycle. Surprisingly, taxi seems to have the highest index in some point which might happen if the person used taxi for a certain trip to a certain destination and for certain activity and if the activity set is small, it can happen. By looking into more disaggregated level (discussed earlier), it is found that most of the taxi users are aged and often travel to personal care, doctor, etc. or conduct a business trip. Looking into the TRI distribution, public transport (train, BTM) and others seems to have higher repetition index in temporal dimension for fixed destinations which represents that respondents are more flexible about departure time when they are using car, bicycle or any private mode, however, if they use public transport, obviously they need to maintain the schedule. Moreover, the mode usage over TRI distribution in flexible destinations have almost similar values and are not very differentiable except the ferry usage. The explanation can be the same as the use of taxi in fixed destinations.

Furthermore, the distribution of repetition index over the travel time classes (classes are defined based on the frequency and statistical distribution, and already showed in Table 4) for fixed and flexible destinations are shown in the Figure A 17. It is remarkable that, in both fixed and flexible destinations, shortest travel time has high SRI followed by sequential next classes. In fixed destinations, the differences are higher than flexible. On the other hand, TRI is found other way around as SRI. Except for >120 minutes, shorter travel times have lower TRI and longer travel

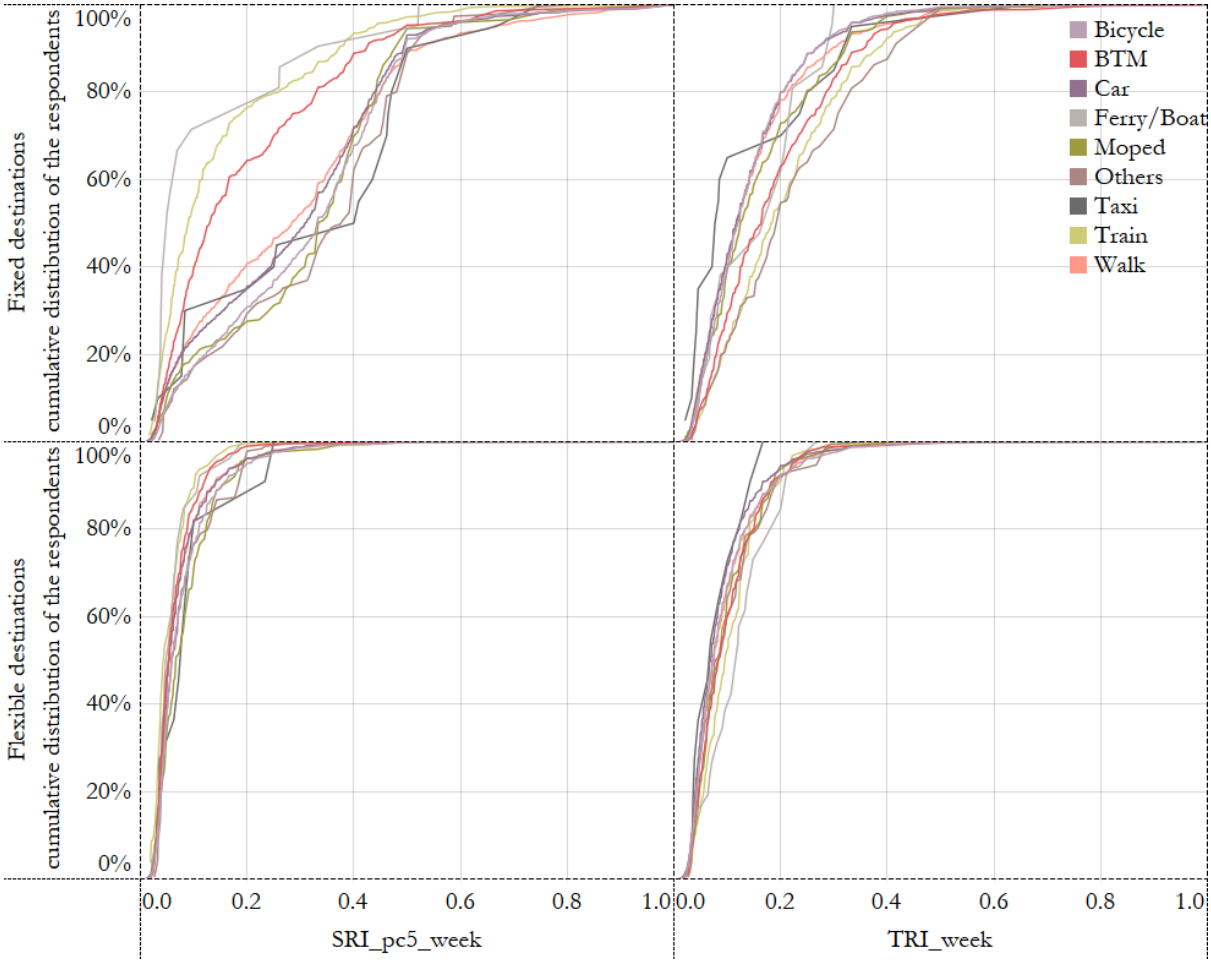

Figure A 16: SRI and TRI distribution over mode Choice

time has higher TRI in fixed destinations. It can be interpreted as, when people have a higher travel time, people are sincerer about time and departs repeatedly at the same time, but shorter travel time means the destination is nearer and therefore, respondents are flexible about departure time. The TRI distribution over flexible destination is difficult to differentiate, except >120 minutes travel time, which seems to have a higher index within (0.15~0.3).

Afterwards, the distribution of the SRI and TRI over departure time is presented in the Figure A 18. The departure time is grouped as the early morning (7-9), late morning (9-12), early afternoon (12-14), late afternoon (14-17), evening (17-20), night (20-24), midnight to dawn (24-4), as discussed earlier. In fixed destination SRI, it can be seen that, early morning has the lowest index than others. While night has the highest SRI, followed by evening, late afternoon, midnight to dawn, early afternoon and late morning. Comparatively high index in fixed destination reflects the high spatiotemporal correlation over flexible. This is because most of the fixed destinations are work, home, appointment, education, which are repeated trips in the same destination. As an exemption, early morning has a very low SRI, which means those trips vary over spatial dimension. On the other hand, the SRI distribution over flexible destinations is sequential, so for instance morning has higher repetition index than the afternoon, followed by evening and night. The TRI of the fixed destination shows that, early morning, late afternoon and evening have higher temporal repetition index than other time classes. This represents the work, study or back home trips. On the other hand, in flexible destination, TRI distribution is dominating in the late morning and late afternoon period.

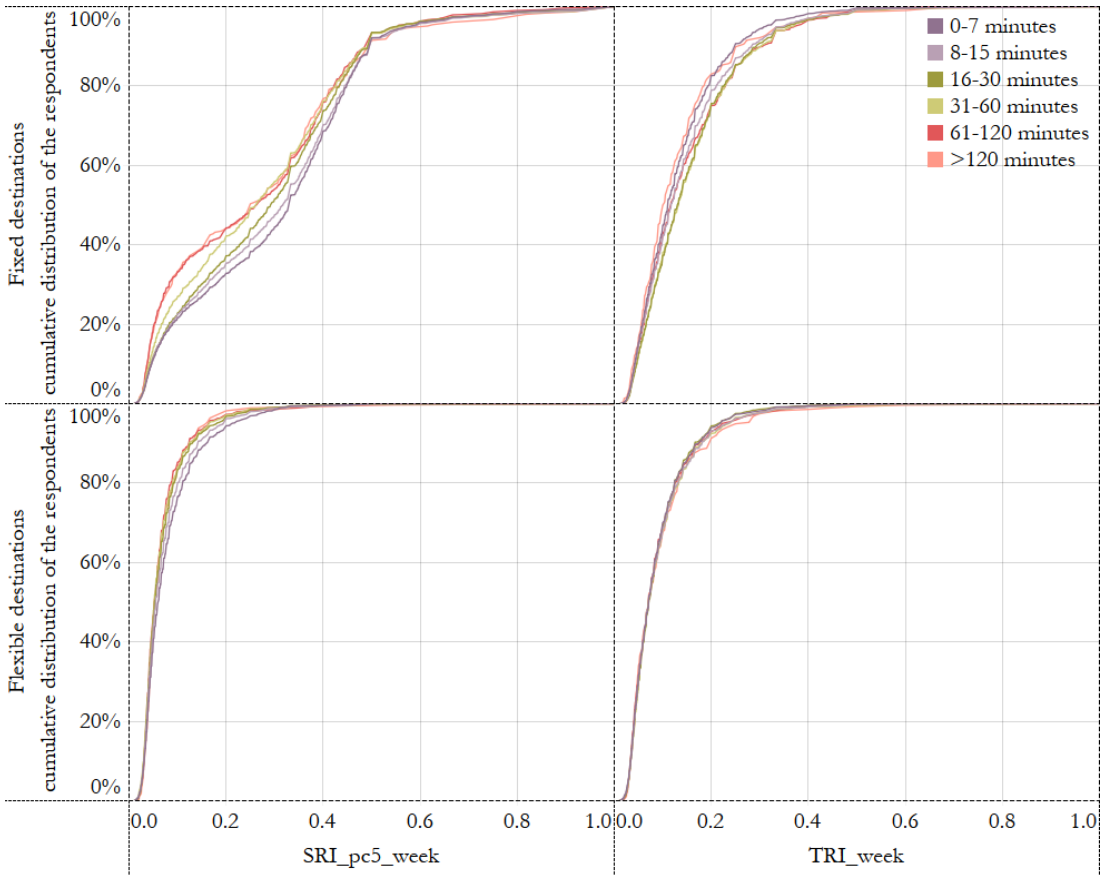

Figure A 17: SRI and TRI distribution over travel time classes

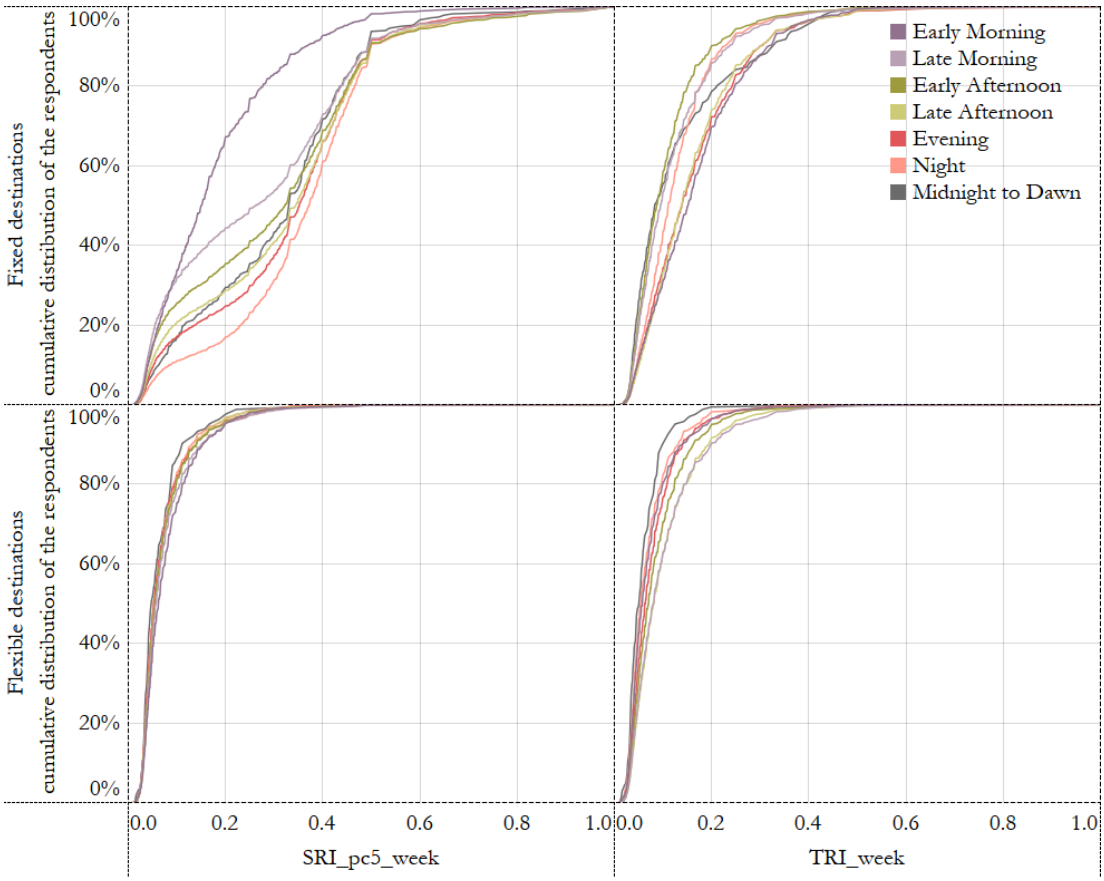

Figure A 18: SRI and TRI distribution over departure time

Others

As Table 7 already showed that the median of SRI in fixed destination is higher during the weekend than workdays and other way around for flexible destination. The cumulative distributed curve is no different than this (Figure A 19) and representing a high intrapersonal variation in destination choice during workdays to fixed destinations and during the weekend to flexible destination. On the other hand, in both destination type, the variation in departure time is higher during the weekend than workdays. This represents that people are more flexible about the departure time during weekends.

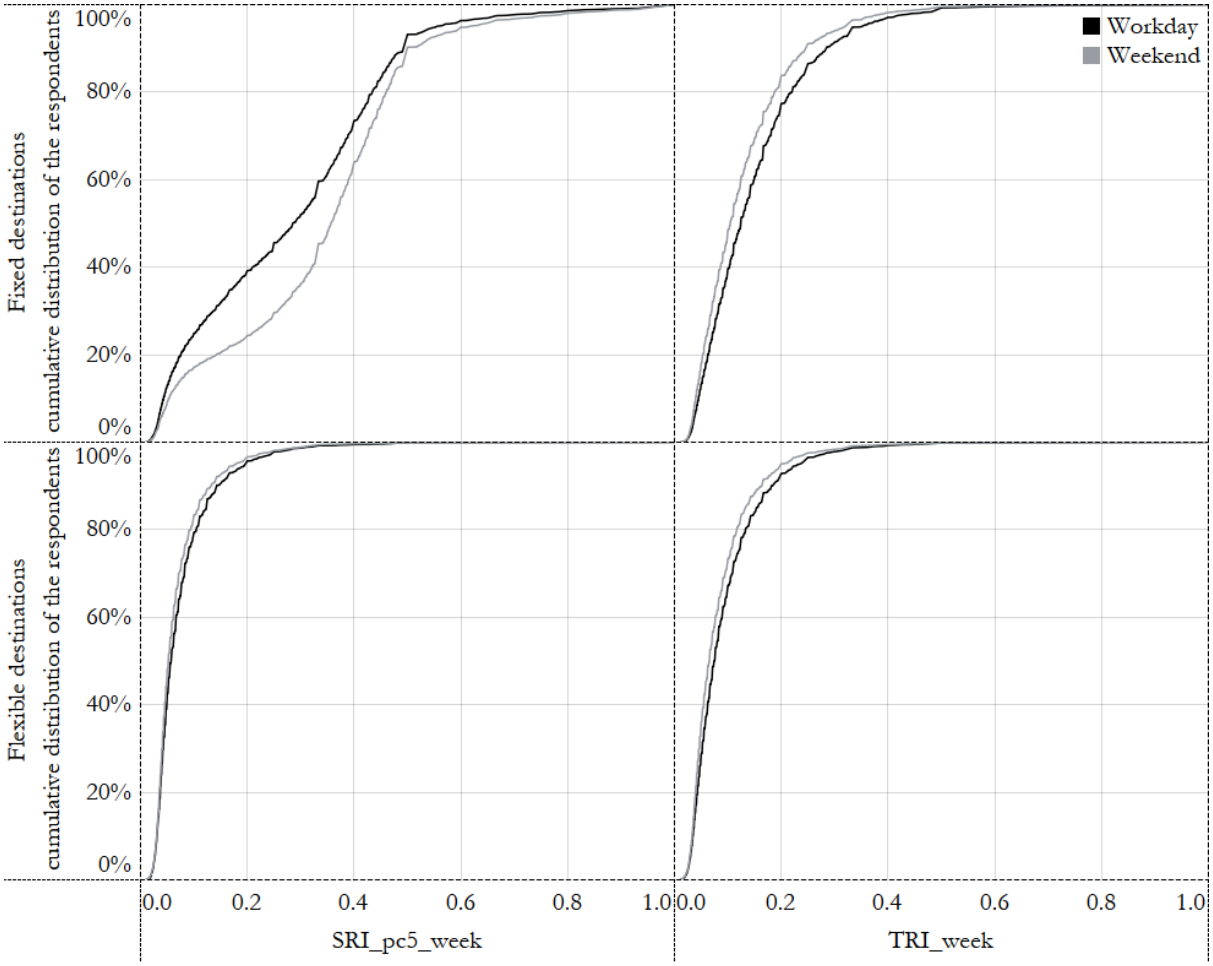

Figure A 19: SRI and TRI distribution over weekdays

There are hardly any differences in the SRI and TRI distribution over weather (Figure A 20). In SRI distribution, the only notable thing is the index value of the cloudy weather is a bit high in fixed destination, means low intrapersonal variation. Remarkable findings have been observed in TRI distribution, although the difference is very small. It can be seen that in both types of destinations, rainy weather has lowest repetition index, which means that the intrapersonal variation in departure time is high when the weather is rainy. This is normal because during the rainy weather, usually people face a dilemma whether they should depart or not. As a result, departure time varies more than cloudy or clear weather.

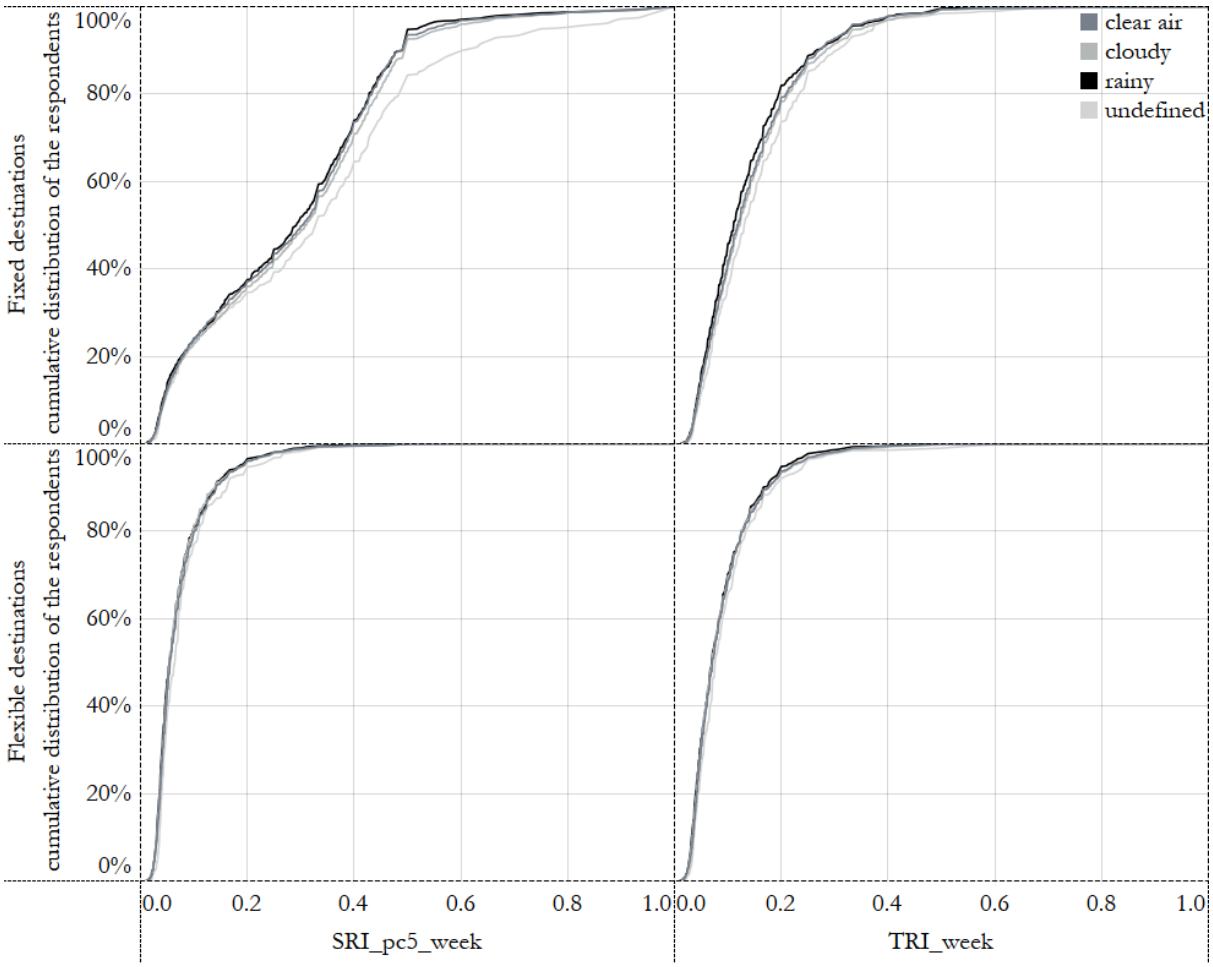

Figure A 20: SRI and TRI distribution over weather

Statistical Tests for Variable Selection

Several statistical tests are carried out to understand the multicollinearity and correlation between the independent and dependent variables. These can affect the model and therefore, correlation analysis is used for checking the correlation, VIF (variance inflation factor) test through linear regression analysis is carried out for checking multicollinearity and dimension reduction (rotated component matrix with principle component analysis) is used to aggregate the correlated similar variables (i.e. spatial variables). SPSS is used for all these tests.

Although these results are not completely reliable because they are simple regression analysis, however to some extent, they can provide insight for the assumption to build the model and help to understand the correlation between variables. The variables having a higher correlation (e.g. correlation coefficient higher than 0.6) can cause problems during estimation process (Simma *et al.*, 2002). For example, the household member has very high correlation with kids in a household variable (0.94), living with partner (0.55), household composition (0.72). Household composition also has a correlation with kids in a household variable (0.66). Further, VIF shows that, household member (83), kids in the household (55), and living with partner (9) seems to have high multicollinearity. Age (33) and age category (25) also have very high multicollinearity (Table A 5). Not to mention, most of these variables are found insignificant in the choice model and therefore excluded from the final model specification. Many spatial variables are found highly correlated with each other. Therefore rotated component matrix (varimax rotation method) with principal components is used to aggregate few variables. For example, farm, grass and orchard are aggregated; heath and nature reserve are aggregated; park, forest and scrub are aggregated; meadow

and vineyard are aggregated; lastly commercial and industrial areas are aggregated together (Table A 6). In the built environment variables, distance to café, cafeteria and restaurant are aggregated as restaurant; distance to cinema, sauna, sunbath, attraction is aggregated as leisure, distance to daycare and after school care are aggregated as daycare (Table A 7). Bike and ride, and walk and ride have a very high correlation (0.95) and consequently they cancel each other's effect in the model (Table A 8).

VIF Test

the variance inflation factor (VIF) quantifies the severity of multicollinearity in an ordinary least squares regression analysis. It provides an index that measures how much the variance (the square of the estimate's standard deviation) of an estimated regression coefficient is increased because of collinearity (*from Wikipedia*). Following threshold is using widely by the statisticians:

VIF	Status of predictors
VIF = 1	Not correlated
1 < VIF < 5	Moderately correlated
VIF > 5 to 10	Highly correlated, high multicollinearity
VIF > 10	Very high multicollinearity

Table A 5: VIF test results

Att	VIF	Att	VIF	att	VIF	att	VIF
Weather	1.008	no_of_jobs_aankompc4	4.822	positie	3.309	woonvorm	2.727
Trip purpose	1.69	SRI_week	2.468	leeftijd	***32.612***	woning	1.318
Mode usage	1.109	TRI_week	1.378	lftdcat	***25.227***	sted	1.453
Mode repetition	1.702	Weekday	1.033	lftdhhh	6.682	belbezig	1.453
Travel time of the trip	2.636	dep_time	1.078	aantalhh	***83.209***	nettoink	1.296
Travel time category of the trip	2.703	Goedgekeurd	1.04	aantalki	***55.23***	netinc	1.22
Population_aankompc4	1.342			partner	9.37	nettocat	1.076
no_of_workers_aankompc4	5.623	Female	1.182	burgstat	2.132	oplcat	1.212

Dimension Reduction

Table A 6: Dimension reduction (rotated component matrix with principle component analysis) of LUT variables

Rotated Component Matrix										
	Component									
	1	2	3	4	5	6	7	8	9	10
Allotments	0.024	0.005	-0.026	-0.042	0	0.018	0.065	0.028	0.847	-0.038
Cemetery	-0.084	-0.022	0.02	0.021	-0.065	0.853	-0.087	-0.018	0.017	0.016
Commercial	-0.081	0.051	0.397	0.154	-0.207	-0.314	-0.254	-0.106	0.336	0.078
Farm	0.647	0.074	-0.174	-0.07	-0.104	-0.04	-0.027	0.019	0.133	0.003
Forest	0.215	0.378	0.419	-0.001	0.042	0.378	0.1	-0.035	0.043	-0.067
Grass	0.744	-0.086	0.194	0.034	0.108	0.057	0.058	-0.03	-0.026	-0.063
Heath	-0.03	0.727	0.012	-0.009	0.34	-0.043	-0.011	0.004	-0.027	-0.002
Industrial	-0.023	-0.033	-0.097	0.967	0.013	0	0.012	0.012	-0.041	-0.029
Meadow	0.044	-0.014	0.23	-0.022	-0.049	-0.099	-0.072	0.633	-0.095	-0.035
Military	-0.02	0.139	0.113	0.027	0.818	-0.078	-0.007	-0.02	-0.053	-0.007
Nature reserve	0.018	0.786	-0.085	-0.006	-0.127	0.014	-0.012	0.011	0.02	0.006
Orchard	0.421	-0.033	-0.164	-0.094	-0.01	-0.117	-0.066	0.042	-0.185	-0.019
Park	-0.024	-0.151	0.656	-0.11	0.127	0.058	0.023	0.084	-0.009	0.004
Quarry	0.02	0.07	0.024	-0.017	-0.168	0.006	0.691	-0.044	-0.12	0.004
Recreation ground	-0.036	-0.075	0.023	0.037	0.13	-0.064	0.675	0.016	0.186	0.011
Residential	-0.745	-0.181	-0.389	-0.451	-0.02	-0.112	-0.045	-0.012	-0.136	-0.152
Retail	-0.019	-0.007	-0.006	-0.017	0.005	-0.001	0.014	-0.001	-0.03	0.992
Scrub	-0.057	0.164	0.37	0.04	-0.365	-0.142	0.114	0.035	-0.231	-0.049
Vineyard	-0.017	0.024	-0.135	0.031	0.022	0.078	0.044	0.782	0.104	0.033

Extraction Method: Principal Component Analysis.						
Rotation Method: Varimax with Kaiser Normalization.						

Based on the analysis, four aggregations are carried out: farm + grass + orchard, heath + nature reserve, park + forest + scrub, meadow + vineyard and commercial + industrial.

Table A 7: Dimension reduction (rotated component matrix with principle component analysis) of BE variables

Rotated Component Matrix	Component			
	1	2	3	4
AF_ARTSPR (doctor)	0.818	0.146	0.206	0.076
AF_SUPERM (super shop)	0.86	0.186	0.121	0.02
AF_DAGLMD (daily grocery)	0.836	0.181	-0.001	-0.028
AF_WARENH (warehouse)	0.372	0.297	0.682	0.105
AF_CAFE (café)	0.578	-0.035	0.425	-0.195
AF_CAFTAR (café plus food)	0.763	0.166	0.08	-0.106
AF_RESTAU (restaurant)	0.753	0.095	-0.03	-0.152
AF_HOTEL (hotel)	0.327	0.297	0.259	-0.227
AF_KDV (day-care for kids/babies)	0.727	0.187	0.258	0.165
AF_BSO (place where students can Stay and do their homework after school hours)	0.744	0.191	0.206	0.19
AF_BRANDW (fire department)	0.551	-0.18	0.241	0.072
AF_OPRITH (onramp of highway)	0.06	0.07	0.058	0.881
AF_TREINST (train station)	0.147	0.77	0.202	0.103
AF_OVERST (transfer point for public transport)	0.135	0.823	0.148	0.002
AF_ZWEMB (swimming pool)	0.261	0.295	0.704	0.026
AF_IJSBAAN (ice skating)	0.126	0.725	-0.028	0.051
AF_BIBLIO (library)	0.634	-0.006	0.4	0.106
AF_BIOS (cinema)	0.112	0.615	0.503	-0.149
AF_SAUNA (sauna)	0.034	0.521	0.353	-0.24
AF_ZONBNK (tanning bed)	0.128	0.505	0.507	0.142
AF_ATTRAC (activities like sport, culture, etc.)	0.042	0.798	0.098	0.047
Extraction Method: Principal Component Analysis.				
Rotation Method: Varimax with Kaiser Normalization.				

Based on the result, AF_LEISURE is defined which is an average of AF_BIOS, AF_SAUNA, AF_ZONBNK and AF_ATTRAC. AF_CAFE, AF_CAFTAR and AF_RESTAU are aggregated and considered the average value as AF_RESTAURENT. Lastly, the average of AF_KDV and AF_BSO are named as AF_DAYCARE.

Correlation Matrix

The correlation matrix for the accessibility measure is presented below. In above cases, dimension reduction and VIF test also consider correlation, therefore not presented anymore. The accessibility measures are highly correlated with each other; therefore, it is expected that one will affects the significance level of another one in the model. Since, BnR and WnR is showing 0.95, mostly they are found cancel each other in the model.

Table A 8: Correlation matrix of job accessibility

	BnR	WnR	Car
BnR	1		
WnR	0.95	1	
Car	0.794	0.686	1

Bivariate Pearson correlation is estimated for trip characteristics and built environment variables and presented below:

Table A 9: Correlation of trip characteristics and built environment

Correlations		EM	LM	EA	LA	E	N	MtD	car	train	btm	moped	bicycle	walk	taxi	ferry	others	TRI_week	mode_rep	BnR	WnR	Car	ARTSPR	SUPERM	OPRITH	TREINST	ZWEMB	IJSBAAN	LEISURE	RESTAURANT	DAYCARE
EM	*Sig. (2-tailed)*		0	0	0	0	0	0	0	0	0	0.35	0	0	0.089	0.68	0.006	0.211	0.072	0	0	0	0	0	0	0	0	0	0	0	0
LM		0		0	0	0	0	0	0.005	0	0	0.004	0	0.629	0.893	0.25	0	0.215	0	0.071	0.033	0.281	0.597	0.597	0.597	0.597	0.597	0.597	0.597	0.597	0.597
EA		0	0		0	0	0	0	0	0	0	0.787	0	0	0.141	0.246	0.43	0.917	0	0.682	0.855	0.397	0.204	0.204	0.204	0.204	0.204	0.204	0.204	0.204	0.204
LA		0	0	0		0	0	0	0	0	0.93	0	0	0.002	0.601	0.67	0	0.168	0.074	0.231	0.318	0.059	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
E		0	0	0	0		0	0	0	0.398	0.174	0.25	0	0	0.14	0.272	0.417	0.517	0	0.37	0.313	0.853	0.307	0.307	0.307	0.307	0.307	0.307	0.307	0.307	0.307
N		0	0	0	0	0		0	0	0.206	0.588	0.349	0	0.276	0.001	0.027	0.704	0.257	0	0	0	0	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
MtD		0	0	0	0	0	0		0.003	0	0.003	0.239	0.085	0.183	0	0.374	0.054	0.129	0.354	0.035	0.048	0.01	0.506	0.506	0.506	0.506	0.506	0.506	0.506	0.506	0.506
car		0	0.005	0	0	0	0	0.003		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
train		0	0	0	0	0.398	0.206	0	0		0	0	0	0	0.415	0.333	0.003	0.516	0	0.104	0.814	0.176	0.719	0.719	0.719	0.719	0.719	0.719	0.719	0.719	0.719
btm		0	0	0	0.93	0.174	0.588	0.003	0	0		0	0	0	0.398	0.315	0.002	0.67	0	0	0.015	0	0	0	0	0	0	0	0	0	0
moped		0.35	0.004	0.787	0	0.25	0.349	0.239	0	0	0		0	0	0.464	0.384	0.008	0.174	0	0.047	0.021	0.357	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49
bicycle		0	0	0	0	0	0	0.085	0	0	0	0		0	0.001	0	0	0.582	0	0.357	0.139	0.573	0.098	0.098	0.098	0.098	0.098	0.098	0.098	0.098	0.098
walk		0	0.629	0	0.002	0	0.276	0.183	0	0	0	0	0		0.033	0.011	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
taxi		0.089	0.893	0.141	0.601	0.14	0.001	0	0	0.415	0.398	0.464	0.001	0.033		0.877	0.64	0.021	0.382	0	0	0	0.577	0.577	0.577	0.577	0.577	0.577	0.577	0.577	0.577
ferry		0.68	0.25	0.246	0.67	0.272	0.027	0.374	0	0.333	0.315	0.384	0	0.011	0.877		0.579	0.426	0.001	0.865	0.944	0.584	0	0	0	0	0	0	0	0	0
others		0.006	0	0.43	0	0.417	0.704	0.054	0	0.003	0.002	0.008	0	0	0.64	0.579		0	0.616	0	0	0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
TRI_week		0.211	0.215	0.917	0.168	0.517	0.257	0.129	0	0.516	0.67	0.174	0.582	0	0.021	0.426	0		0	0	0	0	0	0	0	0	0	0	0	0	0
mode_rep		0.072	0	0	0.074	0	0	0.354	0	0	0	0	0	0	0.382	0.001	0.616	0		0	0	0	0	0	0	0	0	0	0	0	0
BnR		0	0.071	0.682	0.231	0.37	0	0.035	0	0.104	0	0.047	0.357	0	0	0.865	0	0	0		0	0	0	0	0	0	0	0	0	0	0
WnR		0	0.033	0.855	0.318	0.313	0	0.048	0	0.814	0.015	0.021	0.139	0	0	0.944	0	0	0	0		0	0	0	0	0	0	0	0	0	0
Car		0	0.281	0.397	0.059	0.853	0	0.01	0	0.176	0	0.357	0.573	0	0	0.584	0	0	0	0	0		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
ARTSPR		0	0.597	0.204	0.001	0.307	0.001	0.506	0	0.719	0	0.49	0.098	0	0.577	0	0.1	0	0	0	0	0.001		0	0	0	0	0	0	0	0
SUPERM		0	0.597	0.204	0.001	0.307	0.001	0.506	0	0.719	0	0.49	0.098	0	0.577	0	0.1	0	0	0	0	0.001	0		0	0	0	0	0	0	0
OPRITH		0	0.597	0.204	0.001	0.307	0.001	0.506	0	0.719	0	0.49	0.098	0	0.577	0	0.1	0	0	0	0	0.001	0	0		0	0	0	0	0	0
TREINST		0	0.597	0.204	0.001	0.307	0.001	0.506	0	0.719	0	0.49	0.098	0	0.577	0	0.1	0	0	0	0	0.001	0	0	0		0	0	0	0	0
ZWEMB		0	0.597	0.204	0.001	0.307	0.001	0.506	0	0.719	0	0.49	0.098	0	0.577	0	0.1	0	0	0	0	0.001	0	0	0	0		0	0	0	0
IJSBAAN		0	0.597	0.204	0.001	0.307	0.001	0.506	0	0.719	0	0.49	0.098	0	0.577	0	0.1	0	0	0	0	0.001	0	0	0	0	0		0	0	0
LEISURE		0	0.597	0.204	0.001	0.307	0.001	0.506	0	0.719	0	0.49	0.098	0	0.577	0	0.1	0	0	0	0	0.001	0	0	0	0	0	0		0	0
RESTAURANT		0	0.597	0.204	0.001	0.307	0.001	0.506	0	0.719	0	0.49	0.098	0	0.577	0	0.1	0	0	0	0	0.001	0	0	0	0	0	0	0		0
DAYCARE		0	0.597	0.204	0.001	0.307	0.001	0.506	0	0.719	0	0.49	0.098	0	0.577	0	0.1	0	0	0	0	0.001	0	0	0	0	0	0	0	0	

*EM = early morning, LM = late morning, EA = early afternoon, LA = late afternoon, E = evening, N = night, MtD = midnight to dawn

In 90% confidence level, if the value is lower than 0.05, correlation exists, else not. It can be noted that, the variables that have a correlation with another variable, turned out insignificant in the model most of the time.