Technical University of Munich

The Role of Accessibility in Housing Location Choice

Written by:
Fateme Zolfaghari

Submitted to the:
Chair of Modeling Spatial Mobility
Prof. Dr.-Ing. Rolf Moeckel

Advisors:
Prof. Dr. -Ing. Rolf Moeckel
M. Sc. Filippo Contiero (Citilabs, Inc.)
M. Sc. Pedro Donoso (Universidad de Chile)

March 2018

Abstract

Households' location choice is an important aspect of transport planning, which impacts land use, households' travel behavior and demand. Accessibility is widely assumed to be an explanatory factor in defining and explaining households' location choice. The aim of this study is to assess the degree of importance of accessibility in housing location choice besides other parameters for the case study of Richmond, Virginia. Bid-auction approach to modeling location choice and an aggregated logit model were employed to estimate the bid function. US census data and US National Household Travel Survey data were used as the data sources for the study area. Sugar Access (a commercial software solution produced by Citilabs Inc.) and ArcGIS were used to measure three gravity accessibility measurements and two cumulative accessibility measurements. The measurements were made for three destinations: jobs, population and retail jobs. Finally, the enumerated methods were compared with each other. The results of the estimation showed the significance of accessibility in households' location choice in Richmond. Additionally, it was revealed that the traditional cumulative accessibility method outperforms other methods. Also, households value cumulative accessibility measurement to retail jobs more than other destinations.

Technical University of Munich – Professor for Modeling Spatial Mobility

Prof. Dr.-Ing. Rolf Moeckel

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

MASTER'S THESIS

of Fateme Zolfaghari

Date of Issue: 2017-10-01

Date of Submission: 2018-03-30

Topic: The role of accessibility in housing location choice

Accessibility, the ease of reaching the desired destinations, is expected to play an important role for many household types in their location choice. Two groups of significant variables affect the household location choice: the household personal characteristics, such as income, age and educational level, and the location features, like accessibility to different destinations, dwelling alternatives and travel costs.

This thesis aims to provide a comprehensive literature review on determining these factors which affecting housing location choice to implement them for the estimation of a bid function for location choice. Furthermore, the derived function will be examined on the data from the city of Boston, Massachusetts to evaluate the determined factors and to drive the best indices for measuring accessibility. Statistical methods, such as Root Mean Square Error (RMSE) and t-test, will be used to achieve this goal.

The student will present intermediate results to the mentors Prof. Dr.-Ing. Rolf Moeckel and M. Sc. Filippo Contiero in the fifth, tenth, 15th and 20th week.

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

____

Prof. Dr.-Ing. Rolf Moeckel

Declaration concerning the master's Thesis

hereby I confirm that the attached thesis,

" The Role of Accessibility in Housing Location Choice"

was written independently by me without the use of any sources or aids beyond those cited, and all passages and ideas taken from other sources are indicated in the text and given the corresponding citation.

Tools provided by the institute and its staff, such as programs, are also listed. These tools are property of the institute or of the individual staff member. I will not use them for any work beyond the attached thesis or make them available to third parties.

I have not previously submitted this thesis for academic credit.

Munich, March 28th, 2018

F Zolfaghari

Acknowledgments

My greatest appreciation is expressed to Prof. Rolf Moeckel for his encouragement, scientific recommendations and continuous supports.

I wish to acknowledge Citilabs Inc. for providing Sugar Access software. Furthermore, I would like to offer my special thanks to Mr. Filippo Contiero for supplying the data and for his support and feedbacks. I would like to extend my sincere thanks to Mr. Pedro Donoso for his valuable scientific recommendations.

Many thanks to my friend Hema Sheranya Rayaprolu who shared her knowledge and experience with me.

Most special thanks go to my friend Mostafa Fallahnejad, for his encouragement and constructive comments.

Finally, I express my deep gratitude towards my Mum, Dad and family for being there for me all the time and their endless supports.

Contents

List of Figures

List of Tables

Glossary

ABA	Activity Based Accessibility
CBD	Central Business District
GIS	Geographic Information Systems
GTFS	General Transit Feed Specification
ITS	Intelligent Transportation Systems
MNL	Multinomial Logit
NHTS	National Household Travel Survey
NL	Nested Logit
POI	Point of Interest
PT	Public Transport
TSP	Time-Space Prism
UBM	Utility Based Models
USA	United State of America

1. Introduction

Accessibility is highly valued by households in their location choice. In 1973, Wachs et al. defined accessibility as the easiness of reaching destinations for residences (Wachs and Kumagai, 1973). Later in 1999, Wegener et al. completed his definition and pointed to the role of accessibility in their land-use transport feedback cycle (Fig. 1). They indicated that the distribution of infrastructure causes the spatial interactions, these spatial interactions can be measured as accessibility. Furthermore, they stated that the difference in the accessibility of locations not only affects location decisions, but also causes changes in the land use. For example, Improving accessibility of an urban area cause changes in land use and results in dispersed residential development (Wegener and Fürst, 1999). Accordingly, accessibility can be assumed to be an explanatory variable in location choices. However, it should be examined. As an example, this assumption was proved by many independent studies (refer to Table 1).

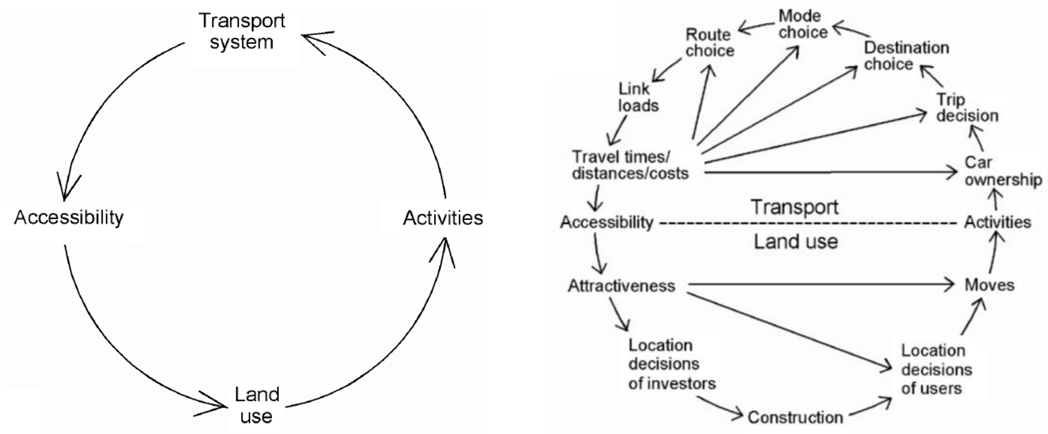

Fig. 1: The land-use transport feedback cycle (Wegener and Fürst, 1999)

Study of housing location choice results in better understanding of households' travel behavior. Glickman et al. showed that households' location choice and short-term decisions such as daily activities are highly integrated (Glickman et al., 2015a). Also, Wegener et al. believe that accessibility of locations has impact on: trip length, trip frequency and mode choice. For example, locations with good accessibility to different destinations make longer trips and more frequent trips. Also, locations with good car accessibility make more car trips and locations with good public transport accessibility make more public transport trips (Wegener and Fürst, 1999). Long-term travel demand can be explained by spatial distribution of households and firms. Their distribution describes trip generation and attraction and can be used to produce origin-destination matrices. Furthermore, spatial distribution in a city can cause many externalities such as congestion, pollution or social segregation. They can also determine land value (Hurtubia and Bierlaire, 2013). Therefore, households' location choice is an important factor in land use modeling, explaining travel behavior, figuring out long-term travel demand and many other factors. Consequently, this study aims to figure out a way to explain and predict households' behavior and the role of accessibility in choosing residing location in Richmond.

There are several accessibility measurement methods that can be measured for different destinations. Srour et al. pointed to the value of determining the best accessibility measurement for defining transportation policies, evaluation of land use and travel network trends. Furthermore, the fact that which destinations are valued more by households leads to better understanding of the land market (Srour et al., 2002).

In this thesis, the different accessibility measurements were addressed in order to determine the most explanatory one in Richmond. The outline of this study is as follows:

Chapter 2: gives a review of different accessibility measurements, their characteristics as well as examples of implementation.

Chapter 3: explains bid-auction approach to modeling housing location choice and methods for estimating it (section 3.1, 3.2, 3.3). Then, the accessibility measurements, data and study area are clarified (section 3.4, 3.5). Subsequently, the expected explanatory variables of the estimation and their correlations are described (section 3.6, 3.7).

Chapter 4: presents the results of estimations and explains the interpretation of the outcomes (section 4.1). furthermore, different accessibility measurements are compared and described (section 4.2, 4.3)

Chapter 5: concludes the report by giving a summary of findings and faced limitations as well as further suggestions (section 5.1).

2. Methods of measuring accessibility

In this chapter, different methods of measuring accessibility and examples of their implementations in housing location choice will be described. Accessibility measurements can be classified into three main categories, location-based, person-based and utility-based. Location-based accessibility measurement analyses accessibility at locations and calculates accessibility to spatially distributes activities, such as jobs and shopping centers. Person-based accessibility measurement analysis accessibility at individual level. Utility-based accessibility measurement analysis the benefits a person can receive from access to spatially distributed activities (Geurs and van Wee, 2004). These three accessibility measurements can be calculated with different methods, these methods were depicted in Fig. 2. Further information regarding these methods, their formulas, examples of application summarized in Table 1 and will be explained in detail later in this chapter. Furthermore, Table 2 summarized the feature of different accessibility measurements. The idea of this categorization was taken from Karst et al. and Geurs et al. publications (Karst and van Eck, 2003; Geurs and van Wee, 2004).

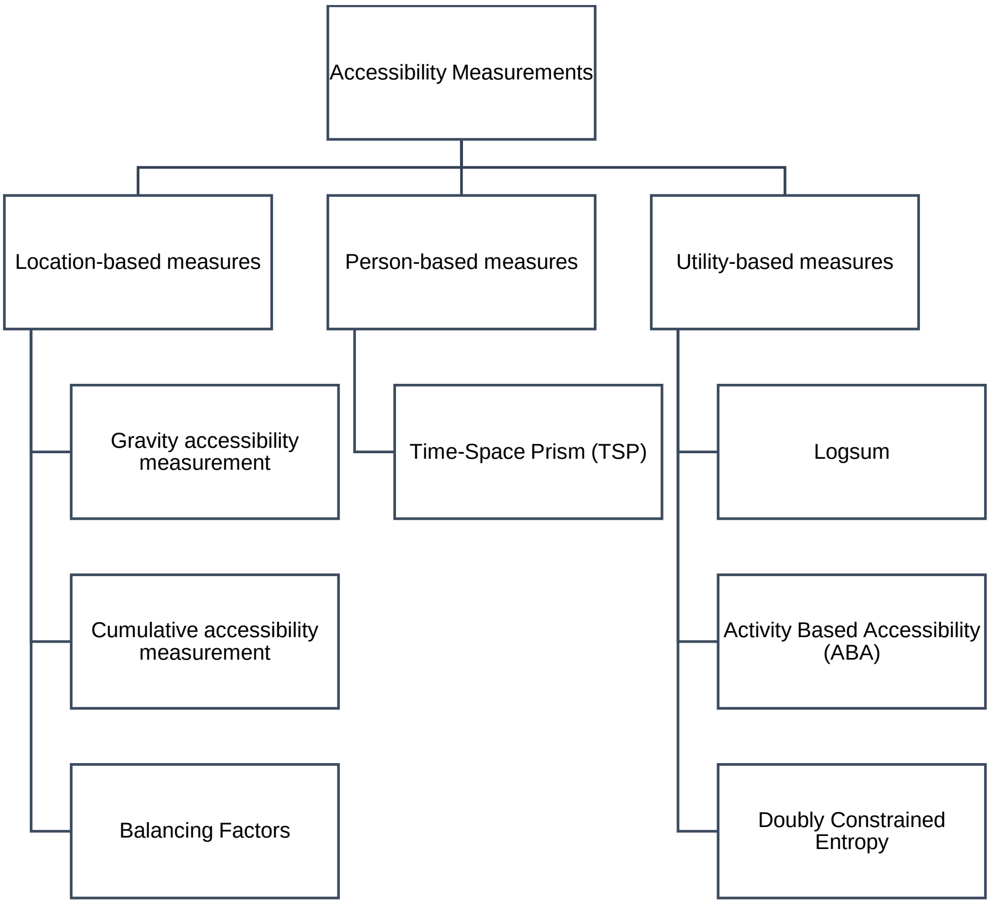

Fig. 2: Some methods of accessibility measurement

Table 1: A summary of different accessibility measurements

Type	Names	Description	Formula	Parameters	Example of application
2.1 Location-based measures	2.1.1 Gravity accessibility measurement	It calculates accessibility of zone i to all other zones (Geurs and van Wee, 2004). Different formulas based on different ways of calculating costs and opportunities have been introduced (Lee et al., 2009b) The summation in formulas is used to show that the calculation considers all potential destinations (j) that might fulfill the desired activity (El-Geneidy and Levinson, 2006a).	$A_{im} = \sum_{j=1} O_j f(C_{ijm})$ $A_{im} = \sum_{j=1} O_j / C_{ijm}$ $A_{im} = \sum_{j=1} O_j \, {C_{ijm}}^{-2}$ $A_{im} = \sum_{j=1} {O_j}^{\alpha} \, exp(\beta C_{ijm})$	A_{im}: accessibility of zone i to possible activity in zone j by using mode m O_j: opportunities at zone j $f(C_{ijm})$: cost function to travel between i and j by using mode m j : each zone of the study area $exp(\beta C_{ijm})$: negative exponential function to travel between i and j by using mode m β: indicates the sensibility of trip maker to C_{ijm}(calibration parameter) α: calibration parameter	▪ **Dallas, Texas -USA** (Guo and Bhat, 2001): Role of school quality and accessibility in residential location choice (UBM) ▪ **San Francisco Bay Area, California-USA** (Guo and Bhat, 2007): The concept of neighborhood and its application to residential location choice (UBM) ▪ **Portland, Oregon-USA**(Dong and Gliebe, 2011)**:** Forecasting the Location of New Housing (UBM) ▪ **Campania-Italia** (Nuzzolo and Coppola, 2007)**:** Relocation of residents and companies as a result of the changes in the accessibility (UBM) ▪ **Buffalo and Seattle Metropolitan Areas-USA** (Hwang and Thill, 2010): The impact of job accessibility on housing prices (hedonic regression modeling) (based on census data) ▪ **Beijing-China** (Wu et al., 2013): The role of job accessibility in residential location choice (UBM)
	2.1.2 Cumulative accessibility measurement	This method calculates the number of opportunity within the predefined range of distance, travel time or cost (Geurs and van Wee, 2004; El-Geneidy and Levinson, 2006b; Lee et al., 2009)	$A_i = \sum_{j=1}^{J} B_j \, a_j$	A_i: accessibility of zone i B_j: a binary value equal to 1, if zone j is within the determined distance (or travel time) a_j: number of opportunities in zone j	▪ **Dallas-USA** (Srour et al., 2002)**:** The role of accessibility in residential location choice and comparing different accessibility measurements (UBM)

Type	Names	Description	Formula	Parameters	Example of application
	2.1.3 Balancing Factors	The balancing factors of doubly constrained spatial interaction model of Wilson can be interpreted as an accessibility measurement (Wilson, 1970; Geurs, 2006). These factors guarantee that the magnitude of flow (like: trips) between zone i and j are equal to the activity numbers in zone i (like: workers) and j (like: jobs).	$a_i = \sum_{j=1}^{n} \frac{1}{b_j} D_j e^{-\beta c_{ij}}$ $\mathrm{b_j} = \sum_{\mathrm{i=1}}^{\mathrm{m}} \frac{1}{\mathrm{a_i}} \mathrm{O_i e^{-\beta c_{ij}}}$	a_i , b_j: Balancing factors O_i: number of opportunities in origin D_j: number of opportunities in destination C_{ij}: the generalized cost of travel β: cost sensitivity parameter	▪ **Netherlands** (Karst and van Eck, 2003)**:** Measuring job accessibility with three different methods (Potential measure, Joseph and Bantock, Balancing factor) and comparing the results
2.2 Person-based measures	2.2.1 Time-Space Prism (TSP)	In 1970, Hägerstrand developed this measurement from his space–time geography, which is based on individuals' point of view. By considering individual's limited budget for time and the possibility of being in a place, it evaluates individuals' ability to take part in different activities over time. This ability can be measured as individual's accessibility (Geurs and van Wee, 2004; Lee et al., 2009).			▪ **Seattle, Washington-USA** (Lee et al., 2009)**:** Operationalizing TSP accessibility in residential location choice (UBM) ▪ **Portland-USA** (Weber, 2003)**:** The importance of accessibility to major employment centers using space-time accessibility measures calculated by GIS
2.3 Utility-based measures	2.3.1 Logsum	Logsum or traditional utility based accessibility measurement calculates accessibility based on the random utility maximization theory and is firstly introduced by Ben-Akiva and Lerman (Ben-Akiva and Lerman, 1977). This theory assumes that people select among the available opportunities in a way to maximize their benefits.	$A_n = E\left[max_{i \in C_n} U_{in}\right]$ $= \frac{1}{\mu} * ln \sum_{i \in C_n} exp(\mu . V_{in})$	A_n: logsum accessibility i: each opportunity n: each person U_{in}: utility of opportunity for individual n V_{in}: the systematic component of utility C_n: choice set (all available opportunities) μ: scale parameter	▪ **Dallas-USA** (Srour et al., 2002)**:** The role of accessibility in Residential Location Choice and comparing different accessibility measurements (UBM) ▪ **Seattle, Washington-USA** (Lee and Waddell, 2010)**:** the role of work accessibility for residential location (UBM) ▪ **Seattle, Washington-USA (Lee and Waddell, 2010):** measures the role of accessibility (UBM) ▪ **Netherlands** (Zondag and Pieters, 2005)**:** Influence of accessibility on residential location choice (UBM) ▪ **Austin, Texas-USA** (Kockelman and Kalmanje, 2003) (Bina et al., 2006)**:** the importance of accessibility to work, freeways, shopping centers, PT and other opportunities in residential location choice (UBM)

Type	Names	Description	Formula	Parameters	Example of application
	2.3.2 Activity Based Accessibility (ABA)	The new approach to measure the utility based accessibility measures, Activity Based Accessibility (ABA), firstly introduced by Ben-Akiva et al. (Ben-Akiva and Bowman, 1998). The main difference of ABA to the traditional way of measuring accessibility (Logsum) is that in the new approach all activities of all individuals during the whole day, plus the impact of trip chains will be considered (Dong et al., 2006).			▪ **Boston, Massachusetts-USA** (Ben-Akiva and Bowman, 1998)**:** measuring accessibility in activity-based travel-demand model ▪ **Portland-USA** (Dong et al., 2006)**:** measuring accessibility in activity-based travel-demand model to demonstrate the impact of peak period toll ▪ **Tel Aviv-Israel**(Glickman et al., 2015a): tested the role of accessibility in housing location (as a long-term decision) (UBM)
	2.3.3 Doubly Constrained Entropy	In 1995 Martinez developed the following measurements from Williams' (1976) integral transport-user benefit measure and balancing factors of Wilson (Williams, 1976; Martinez C., 1995; Geurs and van Wee, 2004). Based on this formula A_i calculates the expected benefits per trip generated which is representor of accessibility and A_j gives attractiveness and expected benefits per trip attracted (Martínez and Araya, 2000).	$a_i = \sum_{j=1}^{n} \frac{1}{b_j} D_j e^{-\beta c_{ij}}$ $b_j = \sum_{i=1}^{m} \frac{1}{a_i} O_i e^{-\beta c_{ij}}$ $A_i = -\frac{1}{\beta} * ln(a_i)$ $A_j = -\frac{1}{\beta} * ln(b_j)$ $A_{ij} = -\frac{1}{\beta} * ln(a_i b_j)$	A_i: benefits per trip generated A_j: benefits per trip attracted A_{ij}: benefits for trip between zones i and j a_i , b_j: balancing factors O_i: number of opportunities in origin D_j: number of opportunities in destination C_{ij}: the generalized cost of travel β: cost sensitivity parameter	▪ **Different metropolitan areas-USA** (Horner, 2004)**:** "Exploring metropolitan accessibility and urban structure"

Table 2: Features of different accessibility measurements

Type	Name	Simple to calculate/ Easy to understand	Competition effect	Individual level	Considers trip chain	All trip purposes	Time dimension/ Scheduling	Valuing closer opportunities
2.1 Location-based measures	2.1.1 Gravity accessibility measurement	+	-	-	-	-	-	+
	2.1.2 Cumulative accessibility measurement	+	-	-	-	-	-	-
	2.1.3 Balancing Factors	-	+	-	-	-	-	+
2.2 Person-based measures	2.2.1 Time-Space Prism (TSP)	-	-	+	+	+	+	?[1]
2.3 Utility-based measures	2.3.1 Logsum	-	-	+	+	-	-	+
	2.3.2 Activity Based Accessibility (ABA)	-	-	+	+	+	+	+
	2.3.3 Doubly Constrained Entropy	-	+	-	-	-	-	+

[1] Is not clear

2.1 Location-based measures

Location-based accessibility measurement analyses accessibility at locations. This measurement calculates accessibility to spatially distributes activities, like number of jobs (Geurs and van Wee, 2004). This method of measurement describes accessibility to different locations with respect to the limited budget (time, cost, distance etc.) (Karst and van Eck, 2003).

2.1.1 Gravity accessibility measurement

This method is also known as Potential, Hansen, Meyer and Miller, Passive and Active accessibility. It firstly introduced by Hansen in 1959 (Eq. 2-2) and is still one of the widely used methods of calculating accessibility as it is simple to calculate and easy to understand. Later in 1967, Wilson improved Hansen findings and introduced (Eq. 2-4) (Wilson, 1967; Schürmann et al., 1997). The idea of this accessibility measurement is that the attraction of a destination increases with size and declines with distance, travel time or cost (Schürmann et al., 1997). It calculates accessibility of zone i to all other zones (Geurs and van Wee, 2004). Different formulas based on different ways of calculating cost function C_{ijm} and opportunities O_j were introduced (El-Geneidy and Levinson, 2006b; Lee et al., 2009):

$$A_{im} = \sum_{j=1} O_j f(C_{ijm}) \qquad \text{Eq. 2-1}$$

$$A_{im} = \sum_{j=1} O_j / C_{ijm} \qquad \text{Eq. 2-2}^{1}$$

$$A_{im} = \sum_{j=1} O_j \ \ {C_{ijm}}^{-2} \qquad \text{Eq. 2-3}$$

$$A_{im} = \sum_{j=1} {O_j}^{\alpha} \ exp(\beta C_{ijm}) \qquad \text{Eq. 2-4}$$

The summation in equations show that the calculation considers all potential destinations (j) that might fulfill the desired activities. For example, when calculating accessibility to shopping centers, all shopping centers of the study area should be considered in the equation (El-Geneidy and Levinson, 2006b).

A_{im}:	Accessibility of zone i to the possible activity in zone j by using mode m
O_j:	Opportunities at zone j
$f(C_{ijm})$:	Cost function to travel between i and j by using mode m
C_{ij}:	Cost to travel between i and j
j :	Each zone of the study area
$exp(\beta C_{ijm})$:	Negative exponential function to travel between i and j by using mode m

[1] Known as Hansen accessibility

β: Calibration parameter that Indicates the sensibility of trip maker to C_{ijm}[1]

α: Calibration parameter

In 1998, Shen pointed to one of the limitation of this accessibility measurement that this measurement cannot be used when competition effect can happen (like: job accessibility) (Shen, 1998). The reason is that this measurement does not consider capacity limitation (every job position can be taken with one person and not more). This can cause mistakes and misleading results (Karst and van Eck, 2003).

Examples of application

In 2001, Guo and Bhat used this method to test the role of school quality and accessibility in residential location choice in Dallas. They found out accessibility to schools is important factor especially for educated households (Guo and Bhat, 2001). In 2007, in their other study, they used the same measurement to calculate the accessibility to three different activity locations shopping, recreational, and employment. For opportunity parameter (O_j), they considered number of retail employment, basic employment and vacant land. Also, the distance between each zone as cost function (Guo and Bhat, 2007). In a different perception of location-based accessibility measurement, Cascetta measured accessibilities out of formula Eq. 2-4 and named them active and passive accessibility. For a specific origin, active accessibility is the cost of getting to other destinations and passive accessibility is the cost of being reached from other origins (Cascetta, 2001). Nuzzolo et al. measured active accessibility based on O_j equal to number of workplaces in zone j and passive accessibility based on O_j equal to the number of residences in zone j (Nuzzolo and Coppola, 2007). Later, Dong et al. used Eq. 2-4 to measure accessibility in their models of forecasting the new housing places in Portland region. They measured the accessibility for two different modes of transport, auto and public transport and for two groups of retail and non-retail employments. They considered O_j as number of jobs and $f(C_{ijm})$ as travel time between zone i and j (Dong and Gliebe, 2011). In a different study in 2010, Hwang and Thill for Buffalo and Seattle Metropolitan Areas and in 2013, Wu et al. for Beijing used Eq. 2-4 to calculate job accessibility (Hwang and Thill, 2010; Wu et al., 2013)

[1] This parameter is negative to show that longer travel times will decrease accessibility (Schürmann et al., 1997)

2.1.2 Cumulative accessibility measurement

This method is also known as Opportunity, Isochrone, Contour measure, Proximity count or Daily accessibility. This measurement is one of the basic methods of accessibility measurement (Wachs and Kumagai, 1973). This method is called cumulative opportunity as it calculates the number of opportunity within the predefined range of distance, travel time or cost (Geurs and van Wee, 2004; El-Geneidy and Levinson, 2006b; Lee et al., 2009):

$$A_i = \sum_{j=1}^{J} B_j \, a_j \qquad \text{Eq. 2-5}$$

B_j: A binary value equal to 1, if zone j is within the determined distance, travel time and cost

a_j: Number of opportunities in zone j

This method is easy to understand and simple to calculate. The problem of using this method is that if the distance, travel time or costs of a specific destination is slightly more than the predefined one, the formula will not count that destination. For example, if the predetermined distance is 400 meters and the opportunity placed in 401 meters, that opportunity will not be counted. Furthermore, this method does not value closer opportunities more than further opportunities. For example, the opportunity which is in 2 meters will be behaved similar to the opportunity that is located in 399 meters, so the method is in the danger of false prediction. Introducing weights for each destination solves this problem (this method will be explained and measured in this thesis). In addition, this measurement not only ignores the competition effect but also does not consider any assumptions on personal preferences of the users and their perception of the transport system and land use (Geurs and van Wee, 2004).

Examples of application

Srour et al. implemented cumulative opportunities in their discrete choice model to measure the role of accessibility to different destinations in housing location chioce. They measured accessibility to three different type of activities: shopping, recreational, and work. To calculate the attractiveness of this destinations, they took into account number of parking spaces and total number of jobs in three different categories, basic employment, retail employment, service employment (Srour et al., 2002).

2.1.3 Balancing Factors

The balancing factors of doubly constrained spatial interaction model of Wilson can be interpreted as an accessibility measurement (Wilson, 1970; Geurs, 2006). The balancing factors are:

$$a_i = \sum_{j=1}^{n} \frac{1}{b_j} D_j e^{-\beta c_{ij}} \qquad \text{Eq. 2-6}$$

$$b_j = \sum_{i=1}^{m} \frac{1}{a_i} O_i e^{-\beta c_{ij}} \qquad \text{Eq. 2-7}$$

a_i , b_j: Balancing factors

O_i: Number of opportunities in origin

D_j: Number of opportunities in destination

C_{ij}: The generalized cost of travel

β: Cost sensitivity parameter

These factors guarantee that the magnitude of flow (like: trips) between zone i and j are equal to the activity numbers in zone i (like: workers) and j (like: jobs). The well accessible places are expected to have a_i smaller than 1, as the number of trips attracted are expected to be equal to the number of opportunities (Geurs, 2006). Balancing factors are useful when competition effect happens in both origin and destination, for example in job accessibility workers are competing for jobs with each other and employers are competing for employees. The iterative process of measuring this accessibility makes it complex and time consuming to be calculated, that could be the reason why this measurement is rarely used (Martínez and Araya, 2000; Karst and van Eck, 2003; Geurs and van Wee, 2004).

Examples of application

In 2003, Karst and van Eck implemented balancing factors in Netherland to measure job accessibility. In the next step of their study, they compared the result of balancing factors with potential accessibility measurement (Karst and van Eck, 2003).

2.2 Person-based measures

This method analyses accessibility at the individual level, for example the possibility of participating in an specific activity with respect to the limited individual's time budget (Geurs and van Wee, 2004). Time space prism is a method to measure this accessibility.

2.2.1 Time-Space Prism (TSP)

In 1970, Hägerstrand developed this measurement from his space–time geography, which is based on individuals' point of view. By considering individual's limited budget of time and the possibility of being in a place, it evaluates individuals' ability to take part in different activities over time. This ability gives individual's accessibility (Geurs and van Wee, 2004; Lee et al., 2009). Weber believes TSP helps planners to have better understanding of urban form by representing individual's accessibility to opportunities in disaggregated level (Weber, 2003). Time-Space Prism (TSP) shows individual's travel pattern, location in time and places a person can reach with respect to the limited time budget. As an example of TSP, Fig.2 shows the person at t_1 is at home and at t_2 at work (Geurs and van Wee, 2004; Lee et al., 2009).

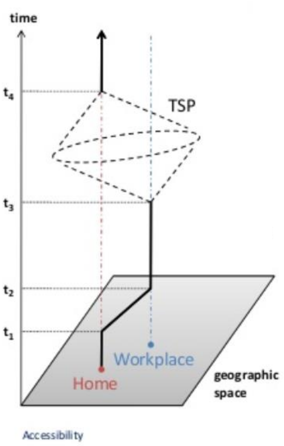

Fig. 2: Time-space prism (TSP) and geographic space (Brian H., 2009)

Examples of application

In 2003, Weber used TSP to measure accessibility in Portland (Weber, 2003). In the next try, in 2009 Lee et al. used this measurement in their utility-based model of residential location. They considered trip-chains in their study and found out non-work accessibility, which was calculated by TSP in a trip-chain setting, will affect housing location decision even after considering accessibility to work (Lee et al., 2009).

2.3 Utility-based measures

This type of measurement focuses on the benefits people achieve by choosing one of the available alternatives, which all satisfy the same need (Karst and van Eck, 2003; Geurs and van Wee, 2004). There are three ways to calculate this accessibility:

2.3.1 Logsum

Logsum or traditional utility-based measurement calculates the accessibility based on the random utility maximization theory and is firstly introduced by Ben-Akiva and Lerman (Ben-Akiva and Lerman, 1977). This theory assumes that people select among the available opportunities in a way to maximize their benefits. The utility of available opportunities (U_i) is not always a known variable and is subjective so it will be considered as a random variable. U_i consists of two components: the systematic utility component (V_i) and the random component (ε_i). The systematic component (V_i) refers to the user's characteristics and the attribute of available opportunities, which both are assumed to affect the decision. The random component (ε_i) represents the unobserved part of the utility (Dong et al., 2006):

$$U_i = V_i + \varepsilon_i \qquad \text{Eq. 2-8}$$

U_i: Utility of opportunity

V_i: Systematic utility

ε_i: Random component

By assuming Gumbel distribution for all opportunities and having the scale parameter of μ, the model will get the form of multinomial logit (MNL). Then, the individua's expected maximum utility can be calculated as:

$$A_n = E\left[max_{i \in C_n} U_{in}\right] = \frac{1}{\mu} * ln \sum_{i \in C_n} exp(\mu . V_{in}) \quad . \qquad \text{Eq. 2-9}$$

A_n: Logsum accessibility

i: Each opportunity

n: Each person

U_{in}: Utility of opportunity for individual n

V_{in}: The systematic component of utility

C_n: Choice set (all available opportunities)

μ: Scale parameter

Examples of application

Srour et al. measured logsum as one of their accessibility measurements for Dallas and compared it with other accessibility measurements (Srour et al., 2002). In Netherland, Zondag et al measured logsum from TIGRIS XL model to implement it as a parameter in their utility-based model. They measured logsum for six types of households and for all possible purposes such as work, education etc. (Zondag and Pieters, 2005). For the city of Austin, Kockelman et al. derived logsum for four different modes of transport, five different time of the day and all available destination combinations (Kockelman and Kalmanje, 2003). In 2009, Bina et al. used the derived logsum of Kockelman et al. to measure the importance of accessibility to work, freeways, shopping centers, public transport (PT) and other opportunities (Bina et al., 2006). In another study, Lee et al. used logsum to measure work accessibility for the home-based trips to evaluate it as an explanatory variable in their utility-based housing location model. Based on available data of two days activities from Puget Sound region council (Seattle, Washington, metropolitan area), they found out that the role of work accessibility for residential location is significant (Lee et al., 2010). In a different study, they used the measured logsum of their previous study to evaluate its importance in their Nested Logit model (NL) (Lee and Waddell, 2010).

2.3.2 Activity Based Accessibility (ABA)

The new approach to measure the utility-based accessibility measurement, Activity Based Accessibility (ABA), firstly introduced by Ben-Akiva et al. (Ben-Akiva and Bowman, 1998). Ben-Akiva et al. used activity-based models and derived accessibility to all activity destinations of individuals. The ABA measures the overall utility of all travel alternatives (Glickman et al., 2015a). Furthermore, the ABA considers impact of trip chaining and the schedule of the activity that differs this accessibility measurement from logsum, which focuses on a particular trip purpose and does not consider trip chaining and time dimension. In other words, ABA considers accessibility to the all activities of all individuals during the whole day (Dong et al., 2006).

Examples of application

The first implementation of ABA was on the activity-based travel-demand model of Boston. Ben-Akiva et al. integrated the demand model with the residential choice model to examine the role of accessibility as an assumed explanatory factor for the individual’s maximum utility model. They examined it for different activity schedules (Ben-Akiva and Bowman, 1998). Later, Dong et al. measured ABA for the activity-based model of Portland. They used ABA to compare it with logsum. They showed the main differences between logsum and ABA that logsum considers only one trip purpose to a specific destination, by one mode of transport at one specific time without taking into account the trip chains. In comparison, ABA covers not only trips to different destinations during the whole time of the day, but also covers travel time and cost of using different modes of travel (Dong et al., 2006). In a different study for Tel Aviv, Glickman et al. proved that accessibility to the main activity destinations is an important factor in housing location choice (Glickman et al., 2015b).

2.3.3 Doubly Constrained Entropy

In 1995 Martinez developed the following measurements from Williams' (1976) integral transport-user benefit measure and balancing factors of Wilson (Williams, 1976; Martinez C., 1995; Geurs and van Wee, 2004):

$$a_i = \sum_{j=1}^{n} \frac{1}{b_j} D_j e^{-\beta c_{ij}} \qquad \text{Eq. 2-10}$$

$$b_j = \sum_{i=1}^{m} \frac{1}{a_i} O_i e^{-\beta c_{ij}} \qquad \text{Eq. 2-11}$$

$$A_i = -\frac{1}{\beta} * ln(a_i) \qquad \text{Eq. 2-12}$$

$$A_j = -\frac{1}{\beta} * ln(b_j) \qquad \text{Eq. 2-13}$$

$$A_{ij} = -\frac{1}{\beta} * ln(a_i b_j) \qquad \text{Eq. 2-14}$$

A_i:	Benefits per trip generated (relative accessibility benefit travelers receive at each origin i)
A_j:	Benefits per trip attracted
A_{ij}:	Benefits for trip between zones i and j
a_i , b_j:	Balancing factors
O_i:	Number of opportunities in origin
D_j:	Number of opportunities in destination
C_{ij}:	The generalized cost of travel
β:	Cost sensitivity parameter

Based on this formula A_i calculates the expected benefits per trip generated which is representor of accessibility and A_j gives attractiveness and expected benefits per trip attracted. In other words, Martínez believes by calculating these formulas researchers can measure the economic benefits of landowners (Martínez and Araya, 2000). The advantage of Doubly Constrained Entropy over Logsum is that it considers the competition effect (Geurs and van Wee, 2004).

Examples of application

Horner measured doubly constrained accessibility of workers with the help of ArcGIS for different metropolitan areas in USA (Horner, 2004).

3. The modelling framework

In this chapter, the bid-auction approach to modelling housing location choice and the methods of estimating bid-function will be described. Then, the dataset preparation for this study will be explained.

3.1 Bid-auction approach

There are two methods to model housing location choice, discrete choice approach or utility-based models and the bid-auction approach. In the discrete choice approach, the consumer evaluates the attributes of each available alternative, such as accessibility to work and dwelling characteristics then selects the location, which maximizes his or her utility (McFadden, 1978). The bid-auction assumes that locations are traded in an auction market, in which the consumer bids his or her willingness to pay for a residential unit (Fig. 3) and later the landlord chooses the best bidder (Alonso, 1964; Hurtubia and Bierlaire, 2013). This thesis focuses on the bid-auction approach therefore, a brief introduction to this approach is given below.

Characteristics of the bid-auction approach (Hurtubia and Bierlaire, 2013; Citilabs, 2014):

- The alternatives are households and households are the bidders.
- The decision maker is the landlord.
- It can be assumed that every located household was the best bidder for that location.
- The household will find the location that provides the highest utility.
- The household bids based on three different attributes: zonal attributes, real estate attributes and household attributes.

Figure removed due to possible copyright infringements

Fig. 3: Customer bids his or her willingness to pay (Citilabs, 2014)

3.2 Methods of bid-auction estimation

Bid-auction is a method to measure households willingness to pay for a location (Hurtubia and Bierlaire, 2013). There are different methods to estimate the bid-auction model such as: Ellickson's method, Martínez's method, Lerman and Kern's method and latent auction.

3.2.1 Ellickson's method

Ellickson proposed his method in 1981, in which he assumed that every located household was the best bidder of that location. Then, he generated a Logit model that is conditional on the location and can be calculated with maximum likelihood. Based on the nature of Logit model, only the relative amount of willingness to pay can be estimated (Ellickson, 1981; Hurtubia and Bierlaire, 2013).

3.2.2 Martínez's method

In 1992, Martínez improved Ellickson's Logit method and proposed an aggregated Logit, which aggregates households into homogeneous groups. By aggregation, the number of alternatives will be reduced to manageable numbers. This approach is useful as, in reality, aggregation of alternatives is common (Martinez, 1992). This method was chosen to be implemented in this thesis.

3.2.3 Lerman and Kern

In 1983, Lerman and Kern proposed a method to solve the Logit problem of generating relative amounts, by producing absolute amounts. Their method requires information on the actual prices and rent paid (Hurtubia and Bierlaire, 2013; Lerman and Kern, 1983).

3.2.4 Latent auction

Lerman and Kern's method requires detailed information about actual prices and paid rents, which is not easy to get. In 2014, Hurtubia and Bierlaire proposed a method to overcome this problem by considering the expected maximum bid of households as a latent variable. Their method requires average or aggregated data on prices and rents (Hurtubia and Bierlaire, 2013).

3.3 Examples of implementing the bid-auction approach

Several land use models implemented the bid-auction approach in modeling the location choice of residents: Cube Land (Citilabs, 2014) RURBAN (Miyamoto and Kitazume, 1989), MUSSA (Martínez, 1996), some levels of UrbanSim (Waddell et al. 2003) and ILUTE (Salvini and Miller, 2005). Most of these models, except Cube Land that uses Martinez's method, used Ellickson's method in their estimations (Hurtubia and Bierlaire, 2013).

3.4 Bid-auction modelling framework

In 1964, Alonso in his book, "Location and land use", introduced real estate as an auction market (Alonso, 1964). In this approach, customers bid their willingness to pay for a dwelling. and later, the landlord chooses the customer with the highest bid. Based on the customer's willingness to maximize his or her utility and considering income constraints, customer's willingness to pay can be calculated as: (Hurtubia and Bierlaire, 2013):(Hurtubia et al., 2010)

$$\max_{x,i} U(x, z_i) \qquad \text{Eq. 3-1}$$

i: Location

x : Vector of continuous goods

z_i: Set of attributes

Total amount spent in goods (x) with price (p) and the price of the location should be less that customer's budget so (Hurtubia and Bierlaire, 2013):

$$s.t. px + r_i \leq I \qquad \text{Eq. 3-2}$$

p: Price

r_i: Price of location i (can be assumed as customer's willingness to pay)

I: Customer's available budget

By putting the constrains Eq. 3-2 in Eq.3-1 and assuming equality in the budget constraint (Hurtubia and Bierlaire, 2013):

$$\max_{i} V(p, I - r_i , z_i) \qquad \text{Eq. 3-3}$$

V: Indirect utility function

By defining $\overline{U}$ as fixed maximum utility level, this formula can be derived (Hurtubia and Bierlaire, 2013):

$$r_i = I - V^{-1}(\overline{U}, p,\ z_i) \qquad \text{Eq. 3-4}$$

Jara-Díaz and Martínez proposed that r_i can be assumed as customer's willingness to pay (Jara-Díaz and Martinez, 1999), so the bid function can be expressed as (Hurtubia and Bierlaire, 2013):

$$B_{hi} = I_h - V_h^{-1}(\overline{U},\ p,\ z_i) \qquad \text{Eq. 3-5}$$

Ellickson proved that the bid in Eq. 3-5, B_{hi} , can be defined as a function of location attributes, $B_h(z_i)$. By adding unobserved utility, the formula will get this form (Ellickson, 1981; Hurtubia and Bierlaire, 2013):

$$\widetilde{B_{hi}} = B_h(z_i) + \varepsilon_h = B_{hi} + \varepsilon_h \qquad \text{Eq. 3-6}$$

ε_h: Unobserved utility

$B_h(z_i)$: Function of location attribute or bid-function

The probability that the household h locate in location i and be the best bidder will be (Hurtubia and Bierlaire, 2013):

$$P_{h/i} = Prob\,\{B_{hi} + \varepsilon_h > B_{h'i} + \varepsilon_{h'}, \forall\, h' \neq h\ \} \qquad \text{Eq. 3-7}$$

McFadden assumed that the error term has Extreme Value distribution so the best bid probability or the probability of household h locate in location i would be (Hurtubia and Bierlaire, 2013; McFadden, 1978):

$$P_{h/i} = \frac{exp(\mu\, B_{hi})}{\sum_{g\in H} exp(\mu\, B_{gi})} \qquad \text{Eq. 3-8}$$

H: All the households who are bidding for location i

It is common in reality to aggregate households to homogenous groups. This aggregation results in reducing number of alternatives into manageable numbers. In 1992, Martinez proposed an aggregated Logit for this purpose. He implemented H_h, total number of households group h , to the Logit model so (Martinez, 1992; Citilabs, 2014):

$$P_{h/vi} = \frac{H_h exp(\mu\, B_{hvi})}{\sum_g H_g exp(\mu\, B_{vi})} \qquad \text{Eq. 3-9}$$

This formula can be rewritten as:

$$P_{h/vi} = \frac{exp(\mu\, B_{hvi} + ln(H_h))}{\sum_g exp(\mu\, B_{vi} + ln(H_g))} \qquad \text{Eq. 3-10}$$

H: Households

V: Real estate

i: Zone

3.4.1 Bid-function

The utility function is a mathematical abstraction that indicates the level of satisfaction or happiness that a household receives by choosing each of the alternatives. The utility function will define the bid function, which can be varied from household to household. For example, some households without children may prefer to locate in a zone with higher accessibility to jobs and other households with children may prefer to locate in zones with higher accessibility to schools. The household's willingness to pay or bid for a location can be derived from the household's utility. It is assumed that households' bids depend on three different attributes, household characteristics, zonal characteristics and real estate characteristics (Fig. 4), so the bid function of household h for real estate v in zone i is (Citilabs, 2014):

Eq. 3-11:

$$B_h = \sum \beta_{Zonal\ Attributes} * Zonal\ Attribues + \sum \beta_{Household\ Attributes} * Household\ Attribues + \sum \beta_{Real\ estate\ Attributes} * Real\ estate\ Attribues$$

Figure removed due to possible copyright infringements

Fig. 4: Bid function (Citilabs, 2014)

3.5 Accessibility measuring

In this section, the accessibility measurements that were measured and estimated in this thesis will be described in detail. Furthermore, Sugar access of Citilabs that is an application for ArcGIS will be introduced. This software and ArcGIS were used to calculate the accessibility measurements for this thesis.

Gravity accessibility measurements are still the most common method of measuring accessibility (El-Geneidy and Levinson, 2006b). Based on the frequent usage of this accessibility measurement in the literature, three kinds of gravity accessibility measurements were measured and estimated. Furthermore, cumulative accessibility measurement is also one of the basic methods of accessibility measurement (Vickerman, 1974, Wachs and Kumagai, 1973). Therefore, two methods of measuring this accessibility were chosen to be measured and estimated, the traditional cumulative accessibility measurement and a weighted cumulative accessibility measurement. These accessibility measurements were chosen based on their importance in the literature and data availability (Table 3).

These accessibility measurements were calculated for three different destinations: jobs, population and retail jobs. These destination were used frequently by other researchers as point of interests (Schürmann et al., 1997; Karst and van Eck, 2003; Hwang and Thill, 2010; Wu et al., 2013; Citilabs, 2014). These accessibility measurements were calculated for four different mode of transport, auto, public transport, bike and walk. They were visualized in Appendix D.

Table 3: summary of accessibility measurements in this thesis

Name	Explanation
Gravity accessibility measurement	Based on Eq. 2-2: $A_{im}= \sum_{j=1} O_j / C_{ijm}$
	Based on Eq. 2-3: $A_{im}= \sum_{j=1} O_j \ \ {C_{ijm}}^{-2}$
	Based on Eq. 2-4: $A_{im}= \sum_{j=1} {O_j}^{\alpha} \ exp(\beta C_{ijm})$
Cumulative accessibility measurement	Traditional cumulative accessibility measurement or Destination Summation. Based on Eq. 3-12: $\mathrm{A_i} = \sum_{\mathrm{j=1}}^{\mathrm{J}} \mathrm{B_j}\ \mathrm{a_j}$
	Weighted cumulative accessibility measurement or access score. Based on Eq. 3-13

3.5.1 Gravity accessibility measurements

Three different methods of measuring gravity accessibility were calculated for this thesis. The travel time was considered as cost, C_{ij}, and the number of jobs, population and retail jobs in each zone was considered as opportunities, O_j. The travel time between each zone by four modes of transport, auto, public transport, bike and walking were calculated. The travel time between each zone was extracted with the help of Sugar Access and ArcGIS. The following equations were used:

$$A_{im} = \sum_{j=1} O_j / C_{ijm} \qquad \text{Eq. 2-2}$$

$$A_{im} = \sum_{j=1} O_j \;\; {C_{ijm}}^{-2} \qquad \text{Eq. 2-3}$$

$$A_{im} = \sum_{j=1} {O_j}^{\alpha} \; exp(\beta C_{ijm}) \qquad \text{Eq. 2-4}$$

In this study, for gravity Eq. 2-4 four different combinations of α and β were measured and estimated. α in the Eq. 2-4 takes amounts equal or greater than 1.0 and is defined to reflect agglomeration effect, that means larger facilities can be disproportionally more attractive than smaller ones. β is a negative parameter that emphasizes on nearby destinations by giving greater weights to them (Schürmann et al., 1997). Table 4 illustrates a summary about the impacts of different calibration parameters.

Table 4: Impact of different α and β in Eq. 2-4 (Moeckel, 2017)

$\boldsymbol{\alpha}$	**Impact of $\boldsymbol{\alpha}$**	$\boldsymbol{\beta}$	**Impact of $\boldsymbol{\beta}$**
1.0	low impact of urban centers	0.3	Low impact of travel time
1.0	low impact of urban centers	0.5	High impact of travel time
1.5	high impact of urban centers	0.3	Low impact of travel time
1.5	high impact of urban centers	0.5	High impact of travel time

3.5.2 Cumulative accessibility measurements

Two different cumulative accessibilities were measured and tested, Destination Summation (the traditional cumulative accessibility measurement) and Access Score (the weighted cumulative accessibility measurement).

3.5.2.1 Traditional cumulative accessibility measurement or Destination Summation

Destination Summation or traditional cumulative accessibility measurement counts the number of opportunities, which are accessible within a predefined travel time. In this thesis, number of jobs, population and retail jobs in each zone were considered as opportunities. Predefined travel times for four different mode of transport were extracted from the decay function (which will be explained later). The Destination Summation accessibility metrics were calculated using this equation (Cirilabs, 2016):

$$D_{ik} = \sum_j S_j f(t_{ij}) \; where \;\; f(t_{ij}) = \begin{Bmatrix} 1 \; if \; t_{ij} \leq T \\ 0 \; if \; t_{ij} > T \end{Bmatrix} \qquad \text{Eq. 3-12}$$

D_{ik}: Destination Summation Accessibility to destination k for zone i

S_j: Number of opportunities in zone j

T: Travel time threshold

k: Number of jobs, population and retail jobs

Decay function[1] clarifies how far people are willing to travel by different modes of transport. For example, according to the Fig. 5, people are willing to travel farther by transit rather than by auto. To determine the travel time threshold for measuring Destination Summation accessibility measurement, the decay factor equal to 50% was used. This decay factor gives 20 minutes travel time by auto, 45 minutes by transit, 15 minutes by bike and 10 minutes by walking. The Destination Summation accessibility measurement for these travel times were calculated with the help of Sugar Access and ArcGIS.

Figure removed due to possible copyright infringements

Fig. 5: Decay function (Citilabs, 2016)

[1] The decay function has been calculated based on U.S. National Household Travel Survey data

One of the limitations of this accessibility measurement is that it considers the same value for closer opportunities and farther opportunities. The Access Score accessibility measurement overcomes this limitation by introducing weights of opportunities.

3.5.2.2 Weighted cumulative accessibility measurement or Access Score

This method solves the problem of traditional cumulative accessibility measurement (Destinations Summation) by considering decay factors and adjusting the associated weight of opportunities. In other words, this method considers a weight for each opportunity based on its distance to the origin. In fact, the weight of each opportunity is derived from the decay function. For example, jobs that are 15 minutes away will be valued more in comparison to the jobs that are 30 minutes away. Fig. 6 shows an example of decay factor.

Figure removed due to possible copyright infringements

Fig. 6: Walk travel time decay factor function

The Access Score equation is as follow:

$$D_{ik} = \sum_j S_j f(t_{ij})\, g(t_{ij}) \; where \;\; f(t_{ij}) = \begin{Bmatrix} 1 \; if \; t_{ij} \le T \\ 0 \; if \; t_{ij} > T \end{Bmatrix} \qquad \text{Eq. 3-13}$$

D_{ik}: Weighted cumulative accessibility measurement to destination k for zone i

k: Jobs, population and retail jobs

T: Travel time threshold

$g(tij)$: Travel Time's decay function

The decay factor for farther destinations tends to zero, so the travel time threshold, T, will not be defined. In other words, the decay function tends to zero for farther destinations and automatically gives zero value to the equation.

3.5.3 Sugar Access

To measure different accessibilities, the Sugar Access software by Citilabs and ArcGIS[1] (Esri, n.d.) were used. Sugar Access is an ArcGIS add-in that calculates accessibility to job opportunities or other points of interests (POI) (Citilabs, 2016).

3.5.3.1 Data sources in Sugar Access

Zonal Data: Each aggregated Census Block[2] (Census Block Group) in the study area was served in Sugar Access as a zone. Demographic information such as number of jobs, population and retail jobs are collected from the US 2010 census data and were included in the Sugar Access data package as ArcGIS polygon feature class.

Roadway and Transit Network: for roadway network Sugar uses the network and travel time published by Here navigation company. Here company not only provides a comprehensive and detailed geometric information but also publishes real travel speeds throughout the day (morning and evening peak hour, off-peak hour). For transit network, it captures transit attributes: mode, headway, run times, transit stops and route path from the local GTFS file.

POI Dataset: Here navigation company provides the database of local points of interest.

3.5.3.2 Travel time analysis

The travel times between all pair of zones in the study area were calculated by Sugar Access to generate travel times' matrices. This was done for four different modes of transport: auto, transit, bike and walking. To avoid errors due to the size of the origin and destination zones, travel times were calculated to the centroid of all zones. Total travel times between origin and destination zones consist of three different travel times:

1. Network Egress Travel Time: Time it takes to move across the origin zonal connectors and get to the roadway network.
2. Network Travel Time: Time it takes to move across the roadway network between origin zone connector and destination zone connector.
3. Network Ingress Travel Time: Time it takes to move across the destination zonal connector and reach the destination's centroid.

Travel time for different mode of transport is based on:

[1] Version 10.5.1
[2] Refer to section 3.5.1

For auto: To calculate travel time for autos, the realistic roadway derived travel times will be used. Different time of day like, morning peak, evening peak, afternoon and nighttime can be specified to specify the effect of congestion on the roadway network throughout the day.

For transit: Walk to Transit Travel Time (A)– Time that takes to walk from origin zone centroid to the transit stop of the best route

Transit Wait Time (B) – By considering maximum waiting time, the Wait Time will be calculated relative to transit line's headway
Transit Run Time (C) – Run time as defined by transit line's attribute between origin and destination
Transfer Wait Time (D) – By considering maximum transfer time, the Transfer Wait Time will be calculated relative to transit line's headway
Walk to Destination Travel Time (E) – Travel time to walk from alighting transit stop to the destination zone centroid

For pedestrian and bicycle: travel time of a straight forward line to the centroid of each zone will be used to calculate travel time for pedestrian and bicycle. In this study, Speed of 5 km/h for pedestrian and 15.5 km/h for bicycle were considered.

3.6 Data and study area

The estimations were done for residential market of Richmond, Virginia and its counties. Two main data were available, 2010 US census database and 2009 US National Household Travel Survey (NHTS). Following, there are explanations about these two data sources.

3.6.1 Census data

U.S. census takes place every 10 years and counts every resident in the United States. In 2010, about 74 percent of households in the US returned their census forms by mail. This amount of participation makes it, a highest movement participation in the history of US. The rest of 74 percent households, were counted by other methods (Census, 2010). The released 2010 census data was used in the census block group level, which is defined as the zonal level for this thesis. Census data was used to prepare the household's dataset and as a dataset for calculating accessibility in Sugar Access. (an overview of 2010 census data can be found in Appendix A)

3.6.2 National Household Travel Survey (NHTS)

The 2009 NHTS was carried out over a period from March 2008 through May 2009. The survey contents five groups of information from each surveyed household: (1) Households (2) Personal (3) Vehicles (4) Worker and (5) Daily travel data. The required data for creating this study's dataset was extracted from group (1) Households and (2) Personal. Table 5 shows an overview of NHTS (NHTS, 2009).

Table 5: Summary of NHTS data

For Each Household:	For Each Person:
• Number of people, drivers, workers and vehicles • Income • Housing Type • Owned or rented • Number of cell phones • Number of other phones • Race of reference person • Hispanic status of reference person • Tract and Block Group characteristics • Daily Travel Data: • Internet Use & Delivery to households	• Age/Sex/Relation to reference person • Driver status • Worker status/Primary activity • Internet use • Home deliveries from Internet shopping • Travel Disability • Effect of disability on mobility • Education level • Immigrant status • Annual miles driven • Incidence of public transit use in past month • Incidence of motorcycle use in last month • Incidence of walk and bike trips in past week • School travel (children)

3.6.3 Dataset of study area

To create a dataset for this thesis (example in Appendix B), each census block group of the census data was considered as a zone. In total, there were 746 zones in the study area, these zones were visualized[1] in Fig. 7, Fig. 8 and Fig. 9 By using reported longitudes and altitudes of NHTS data, the positions of the surveyed households were imported to the ArcGIS (Fig. 10). Three different attributes: (1) Zonal Attributes, (2) Household Attributes (3) Real Estate Attribute were extracted from the data sources to create the dataset. Out of 2,187 interviewed households in the study area, 2,010 household were valid and were included in the dataset. The other 177 households were deleted due to their confidential information of income. Reported weights from the NHTS were also used for the households to minimize biases and adjust the survey. As it stated, accessibility to three opportunities: jobs, population and retail jobs were measured. Fig. 11, Fig. 12 and Fig. 13 visualize the density of these opportunities in each zone. As it was expected, the central zones of Richmond (downtown) have the highest number of all mentioned opportunities.

[1] All visualization was done with ArcGIS 10.5.1

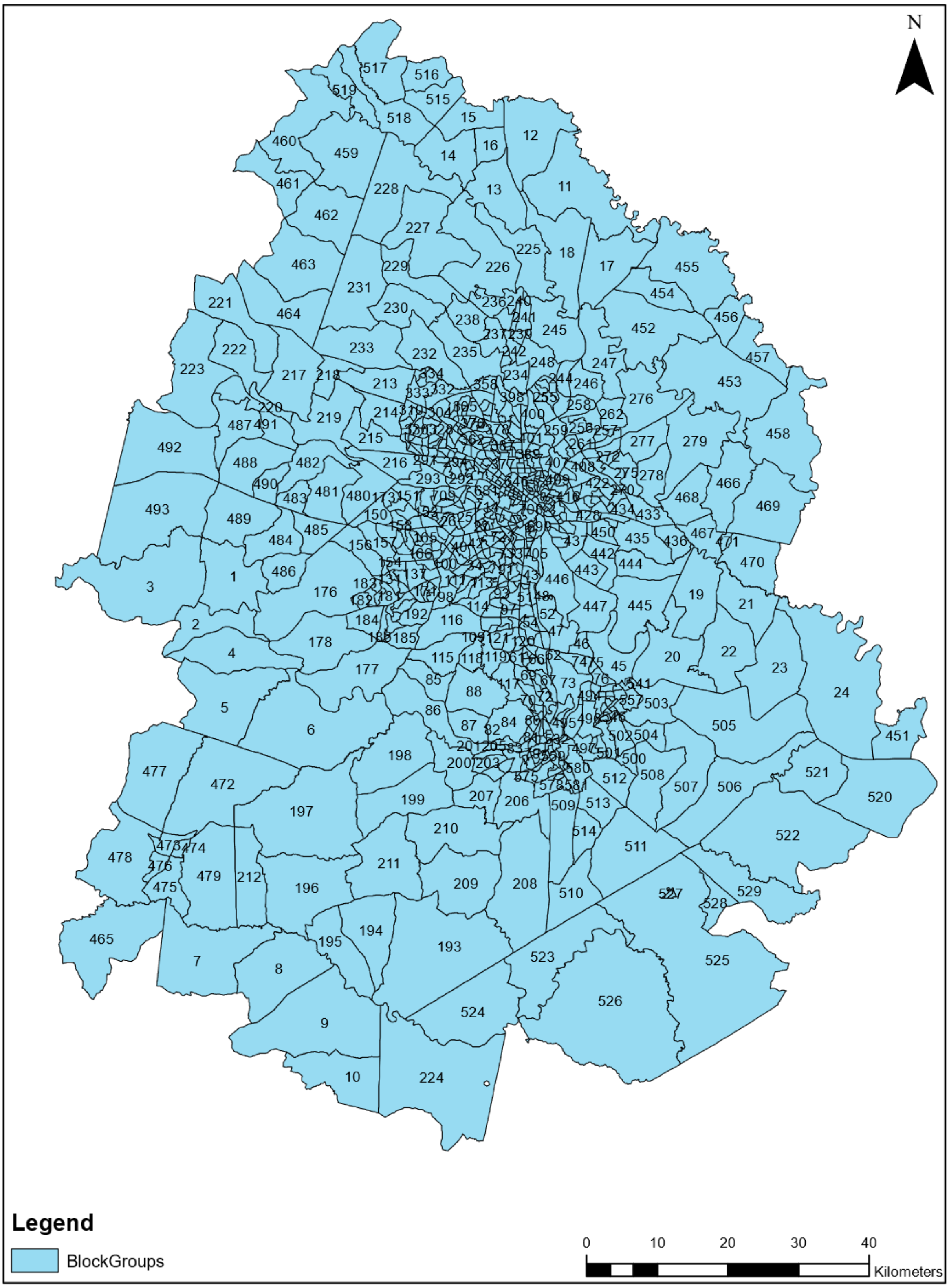

Fig. 7: Zones of the study area

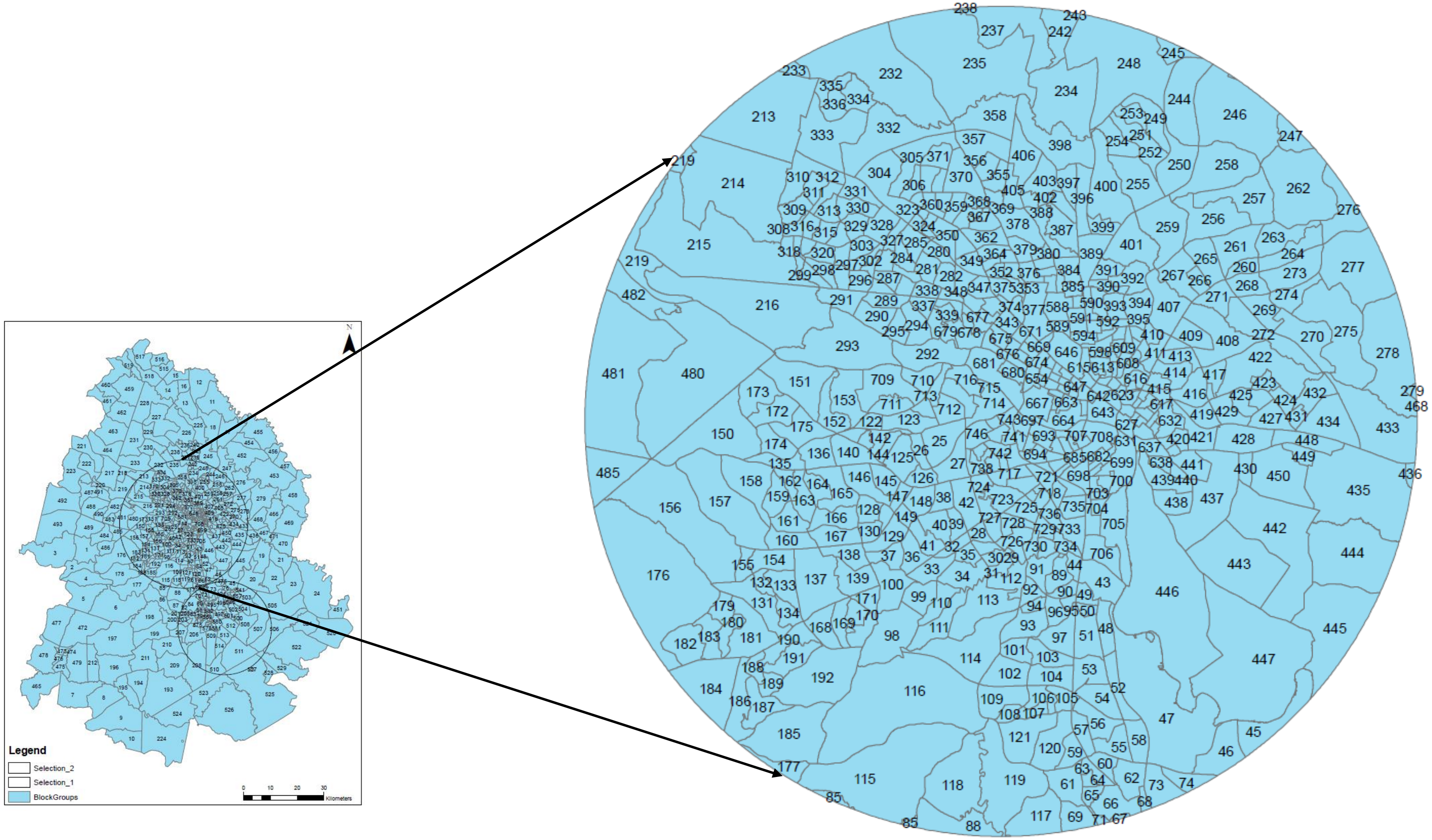

Fig. 8: Zones of the study area (more detailed 1)

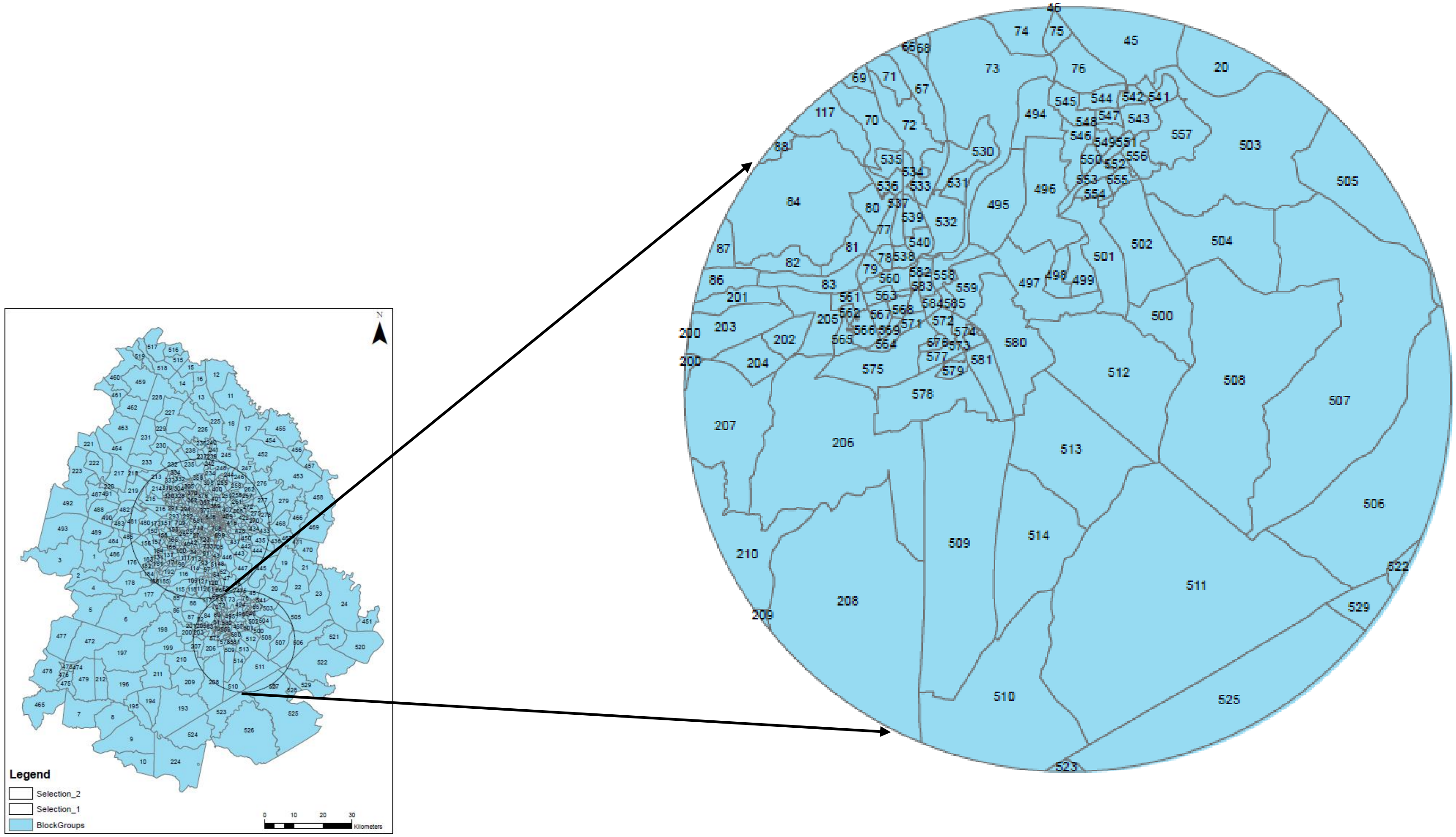

Fig. 9: Zones of the study area (more detailed 2)

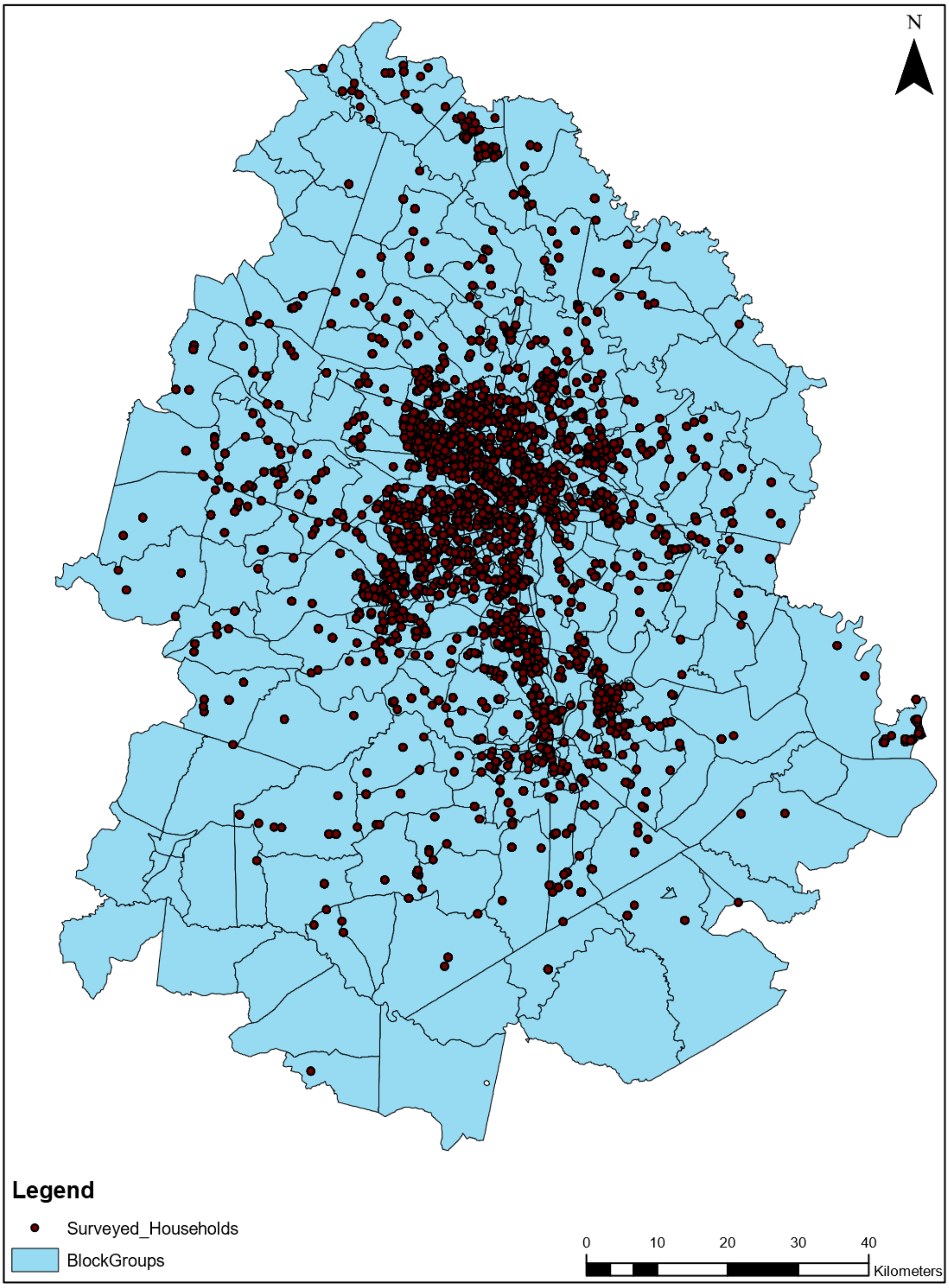

Fig. 10: Place of each surveyed households

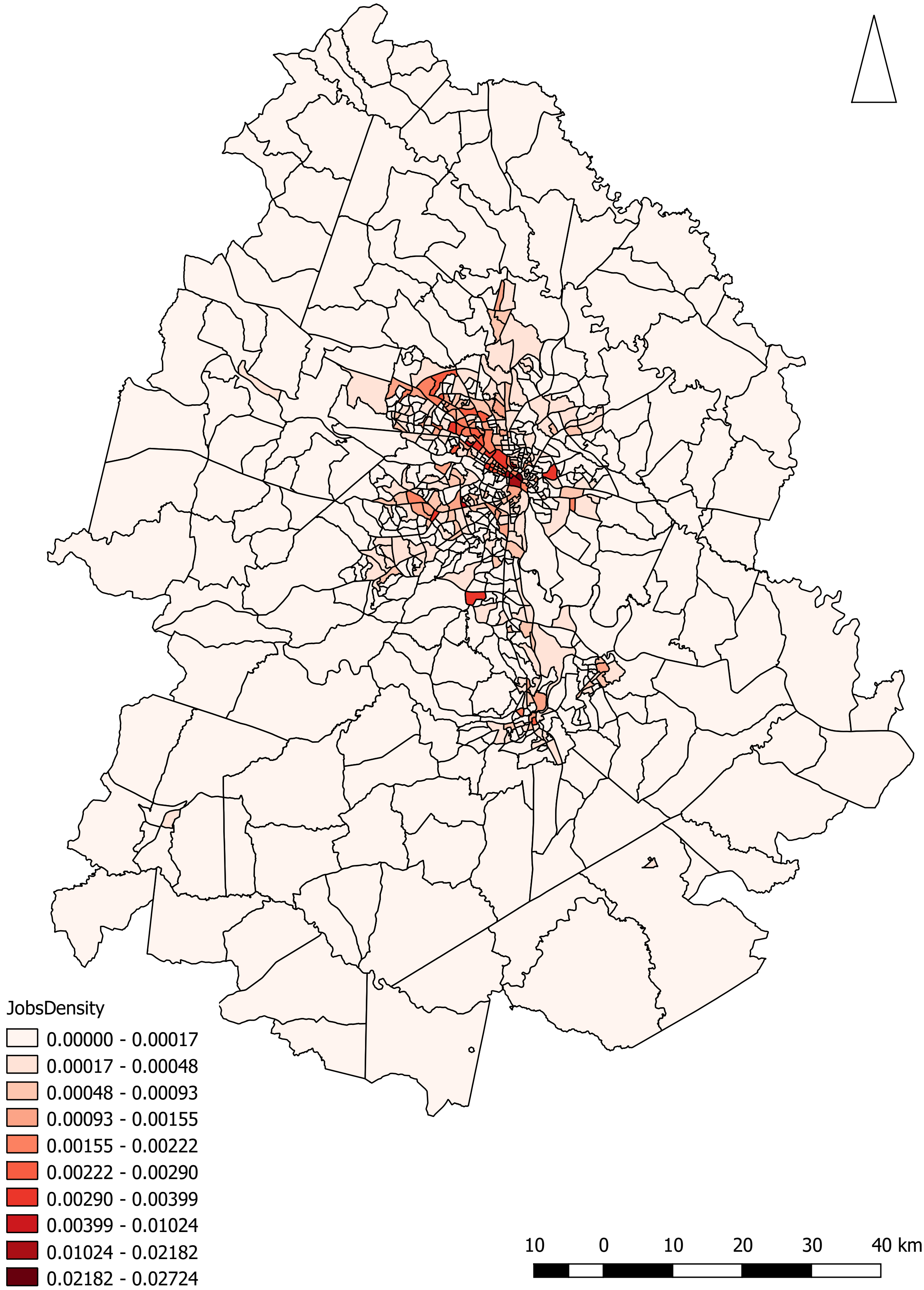

Fig. 11: Jobs density of the study area

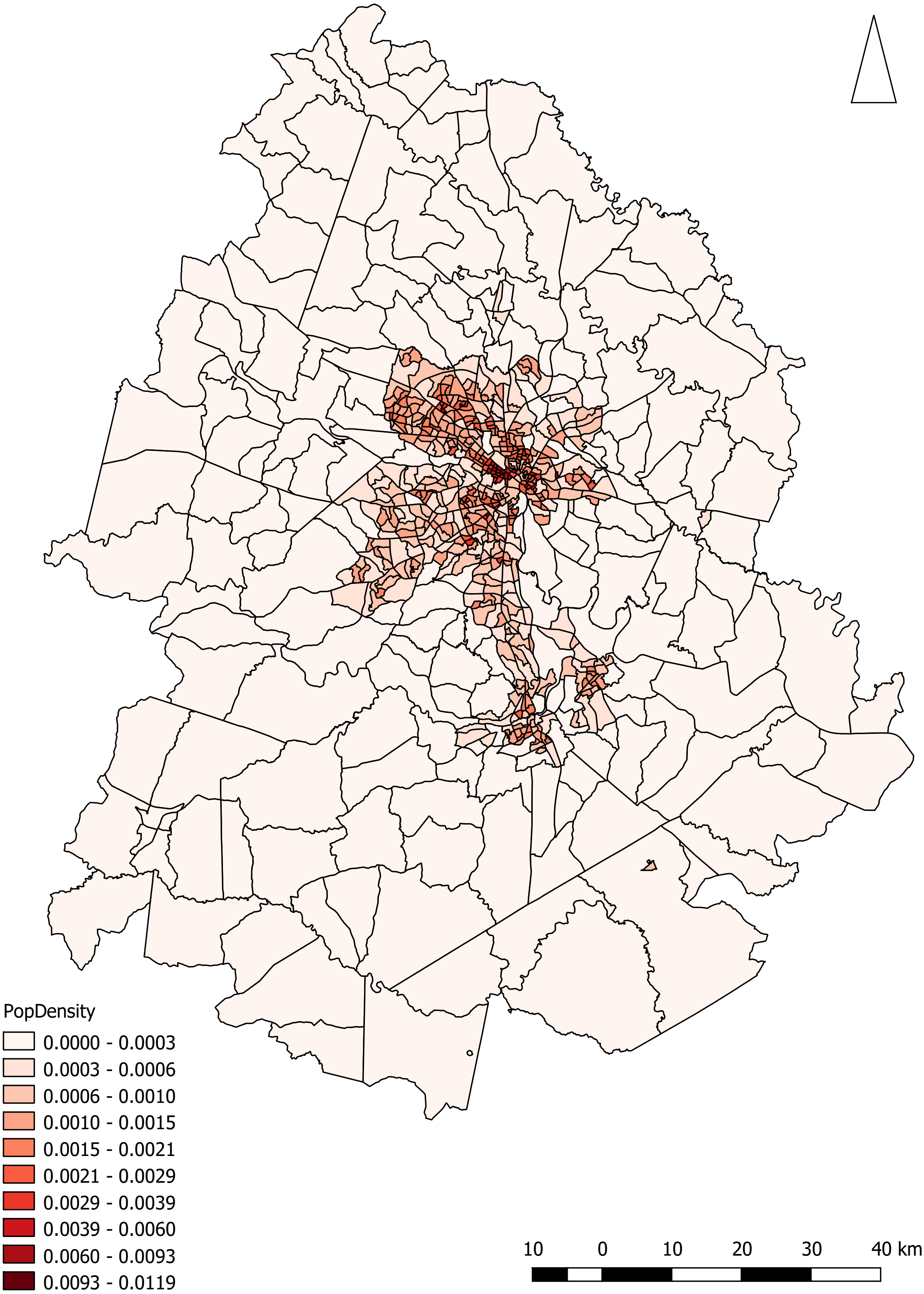

Fig. 12: Population density of the study area

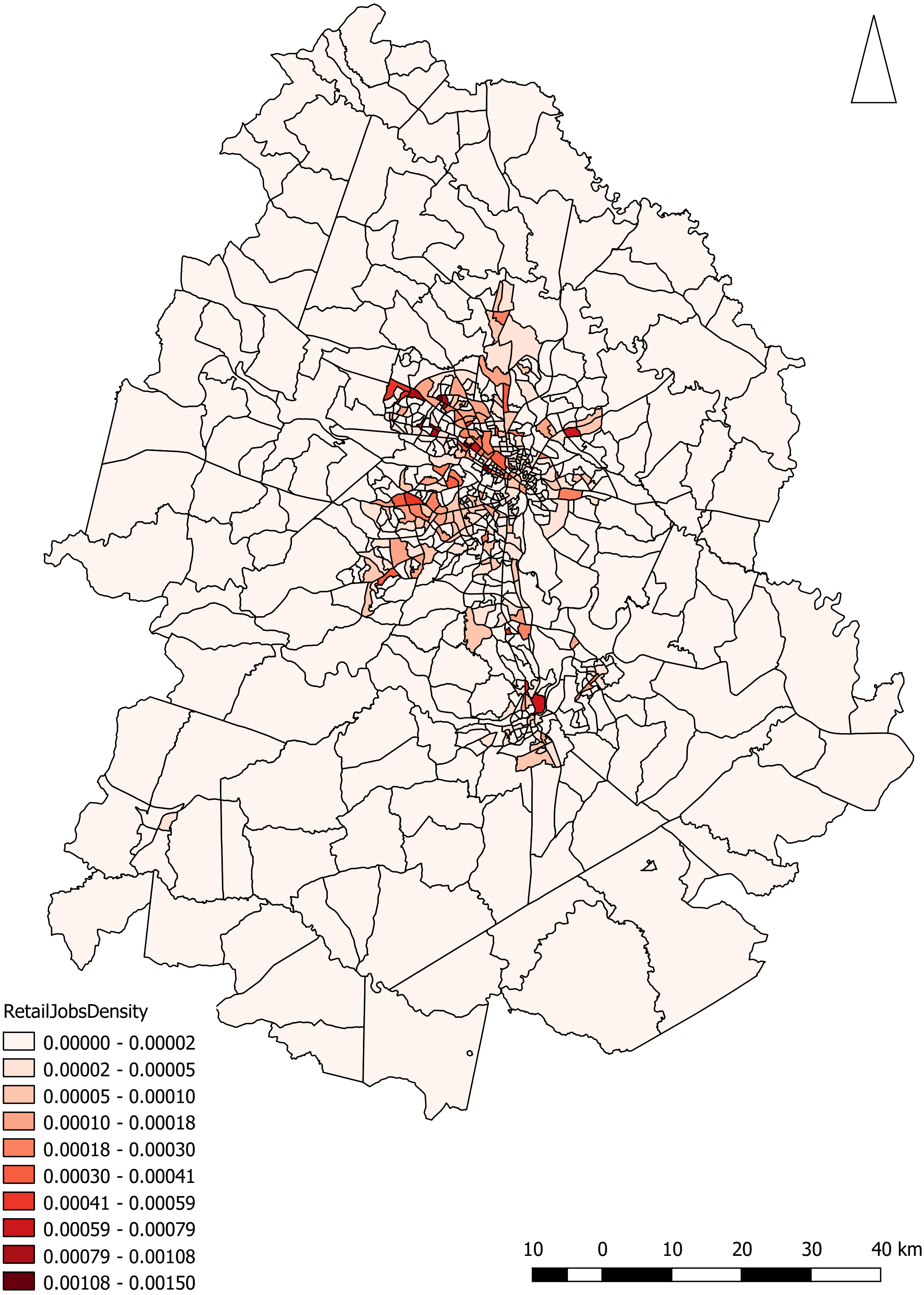

Fig. 13: Retail jobs density of the study area

3.7 Explanatory variables

The explanatory variable where grouped into two groups: independent variables and dependent variables. Later, the coefficient of independent variables will be measured, and their significance will be discussed. Example of other studies explanatory variables can be found in Appendix C.

3.7.1 Independent variables

Table 6 shows a summary of the independent variables considered in the estimations. These variables were chosen based on the available data and their importance in the literature. They categorized to three groups: zonal attributes, household attributes and real estate attributes.

Table 6: Independent Variables considered in the estimation

Independent Variables	Source of the data
Zonal Attributes	
Accessibility	Census and Sugar Access
Urban or Rural	NHTS
Number of schools	Census and Sugar Access
Number of schools * Child	Census and NHTS
% High income * High income	Census and NHTS
% Low income * High income	Census and NHTS
% Educated * Education	Census and NHTS
% Race * Race	Census and NHTS
Zonal income	Census
Household Attributes	
Retired	NHTS
Children	NHTS
Age	NHTS
Gender	NHTS
Education	NHTS
Race	NHTS
Real estate Attributes	
Home type * household size	Census and NHTS

Zonal attributes

Accessibility: Based on three reasons, accessibility measurement of car among other modes of transport: transit, bike and walking, was chosen to be tested. Firstly, as it is shown in Fig.14 for commute trips to work (as an example) 91.5% of the trips are made by car, means it can be assumed most of the trips of households will be made by car. Secondly, most of the households have at least one car and just 2% of them do not have car (Fig. 15). Finally, testing the correlation between accessibilities of different modes of transport shows that these accessibility measurements are highly correlated, so they cannot enter to the model at the same time.

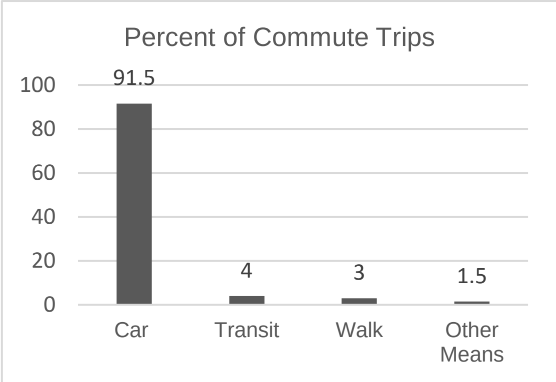

Fig. 14: Percent of commute trips by mode to work (NHTS, 2009)

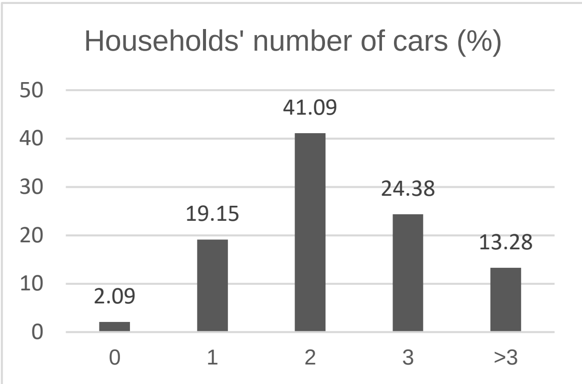

Fig. 15: Auto availability of households in the dataset (NHTS, 2009)

Urban or Rural area: If the dwelling is in an urban area 1 and 0 otherwise. It is expected households prefer to locate in urban areas.

Number of schools: Sugar Access defines school as a POI and measures total number of schools in each zone. A positive value for this attribute is expected so if the number of schools increase, the zone would be more attractive for households.

Number of schools * Child: It is expected households with children are more interested to locate in zones with higher number of schools. Number of schools in each zone was measured with Sugar Access and if the household has at least one child, the child parameter gets 1 and 0 otherwise.

% High income * High income and % Low income * High income: Percentage of high and low-income households of each zone were extracted from the census 2014[1] and income of each surveyed household was available from NHTS. It is expected high income households prefer to be surrounded by high income households, so *% High income * High income* predicted to have positive value and high-income households do not prefer to be surrounded by low income households, so negative value for *% Low income * High income* was expected.

% Educated * Education: Percentage of educated people in each zone was calculated from the census 2013[1] and the *Education* parameter which shows if the household is educated (equal to 1) or not (equal to 0) was available from NHTS.

% Race * Race: It is assumed that the households prefer to locate in zones that their same race is dominant. To test this assumption, percentage of households' race in each zone was measured from the census data. The race parameter will be explained later in the households' attribute.

Zonal income: A dummy variable was defined for this parameter, so if the average income of the zone is higher than average 1 and otherwise 0.

Household attributes

Retired: The purpose of including this parameter was to understand how landlords react to the retirement status of the bidders. A dummy variable was defined to represent this parameter, which takes 1 for households who are retired and 0 otherwise.

Children: A dummy variable was defined to represent the existence of children in the household. It takes 1 if the household has children and 0 otherwise. The aim was to see how presence of children affect the chance of winning the bid auction.

Age and Gender: Age of the oldest member of the household and the respected gender were defined as age and gender of the head of the household. This assumption that the oldest member is the head of the household causes some limitations. For example, if the grandparents are living with the family, although they do not make important decisions of the household like choosing households' location, they will be chosen as the head of the household. This parameter gets 1 if the gender is male and 0 otherwise. The frequency distribution of male and female households in the dataset by their ages is shown in Fig. 16.

[1] Because the income data of 2010 is not available, the data of 2014 was used. It is assumed both income data are similar, and this does not affect the estimations.

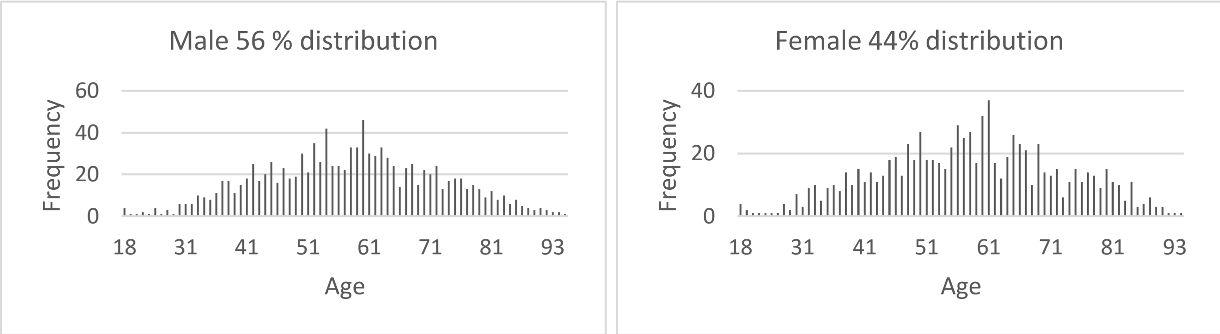

Fig. 16: Distribution of male and female households by age

Education: Highest completed grade of the household is reported by NHTS. A dummy variable was defined that takes 1 for individuals with a college degree and 0 for those without. It is assumed that high educated households are more likely to win the real estate auction.

Race: NHTS reports the race of the households, these races are shown in the first column of Table 7. Census data also reports the number of each race in each zone, this is shown in the second column of Table 7. For households with Multiracial, Hispanic/Mexican and Not ascertained race that there is no report for them in the census data, the reported other race in the census were considered. The frequency of households' race in the study area is shown in Fig.17. This parameter is a dummy which takes 1 if the race of the household is similar with the dominant race in the zone and 0 otherwise.

Table 7: Reported races in census data and NHTS

Races reported in NHTS	Race reported in the census data	Chosen Race
White	White	White
African American, Black	Black	Black
Asian Only	Asian	Asian
American Indian, Alaskan Native	American	American
Native Hawaiian, Other Pacific	Pacific	Pacific
Multiracial	-	Other
Hispanic/Mexican	-	Other
Not ascertained	-	Other
-	Other	-

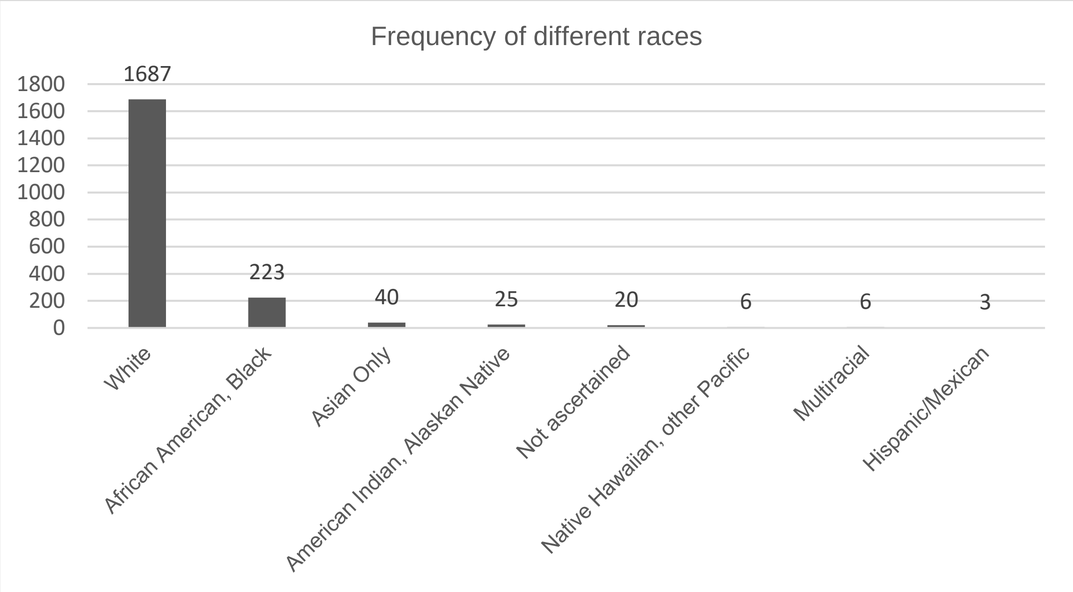

Fig. 17: Frequency of different races in the study area

Real estate attributes

Home type * household size: This parameter was defined to check whether larger households prefer detached houses. Out of 2010 dwellings in the study area 1817 dwellings are detached. The parameter, *household size* which counts the number of households' members was extracted from NHTS.

3.7.2 Dependent variable

Income: Household groups are the dependent variables or alternatives that the landlord chooses the best bidder among them. The NHTS records the total yearly income of the surveyed households and aggregates them into 18 groups as below:

01 =: < $5,000
02 =: $5,000 - $9,999
03 =: $10,000 - $14,999
04 =: $15,000 - $19,999
05 =: $20,000 - $24,999
06 =: $25,000 - $29,999
07 =: $30,000 - $34,999
08 =: $35,000 - $39,999
09 =: $40,000 - $44,999
10 =: $45,000 - $49,999
11 =: $50,000 - $54,999
12 =: $55,000 - $59,999
13 =: $60,000 - $64,999
14 =: $65,000 - $69,999
15 =: $70,000 - $74,999
16 =: $75,000 - $79,999
17 =: $80,000 - $99,999
18 =: > = $100,000

To calculate bid function, households were grouped based on their income. Fig. 18 depicts the frequency of these 18 income groups in Richmond. Each of these 18 aggregated households share bid function coefficient values. Hurtubia et al. and Waddell also aggregated households based on their incomes in their studies (Waddell, 2000; Hurtubia and Bierlaire, 2013).

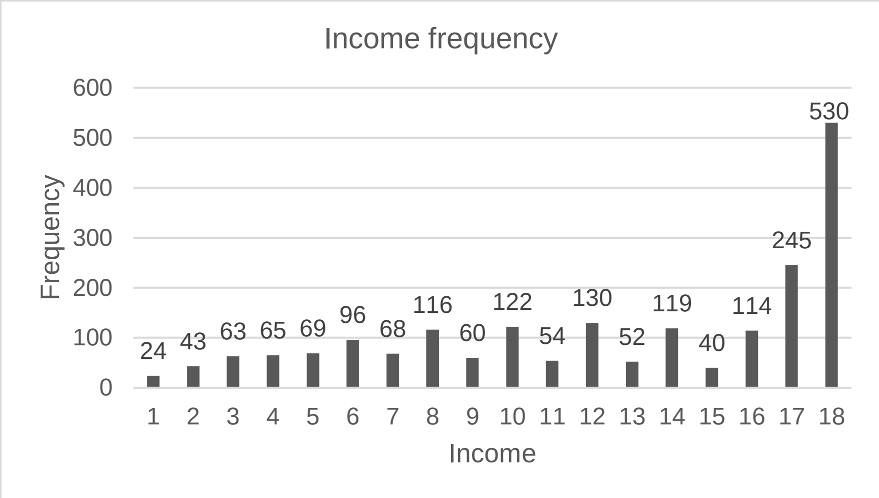

Fig. 18: Frequency of income groups

3.8 Correlation

The correlation between all independent variables were measured to ensure no two variables are highly correlated. To measure and visualize correlations, the 'corrplot' package of R software was used (R, Package 'corrplot,' 2017). R is a free software for statistical computing and graphics.

As it is illustrated in Fig. 19, there is a high correlation between *Number of schools * Child* and *Number of schools*, *Number of schools * Child* and *Children*, *Race* and *%Race * Race*. To avoid the negative effect of these correlations, they were not estimated together in the model at the same time. *Age* was dropped from the estimation as the correlations between *Age* and *Retired*, *Age* and *Children* are high. The correlation shows as the *Age* increases, there is a higher possibility that the person is retired, and it is less likely that the children are still living with the household.

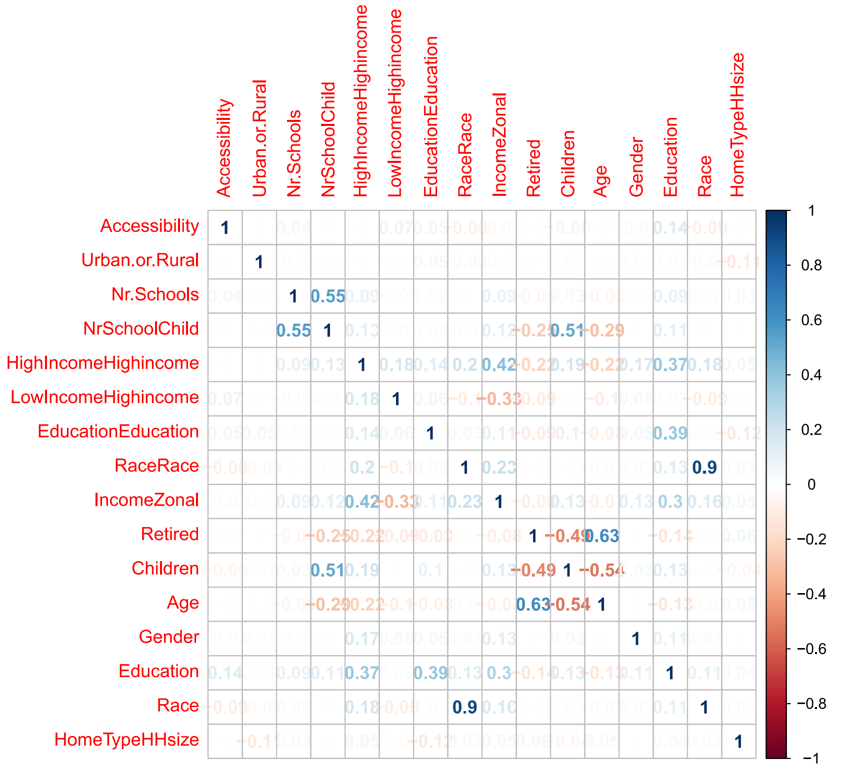

Fig. 19: Variable correlations

3.8.1 Correlation of accessibilities

The correlation between accessibilities of different modes of transport is high (Fig. 20). For example, if a zone has high accessibility to jobs by car, it has also high accessibility to jobs by transit. To avoid negative affect of this correlation, the accessibility by car was entered to the model.

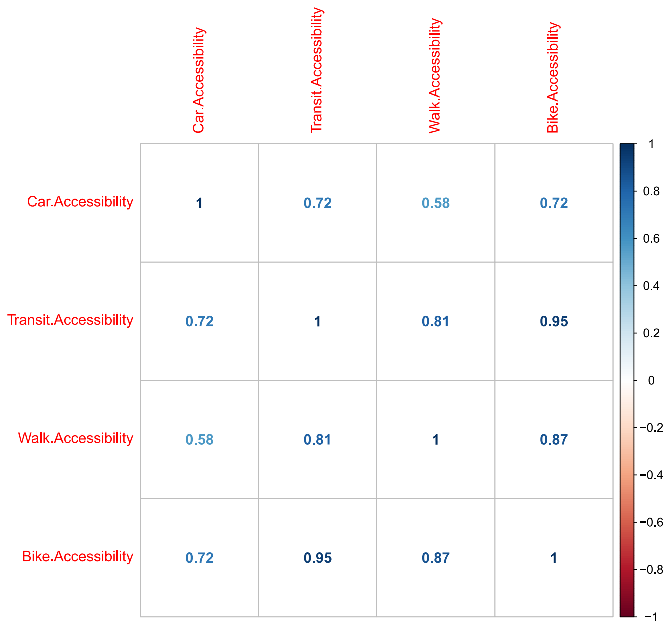

Fig. 20: correlation between accessibility of different modes of transport

Fig. 21 shows correlation between accessibility to different destinations. They are also highly correlated, so they were estimated separately. Srour et al. also mentioned the high correlation between accessibility to different destinations in their study (Srour et al., 2002).

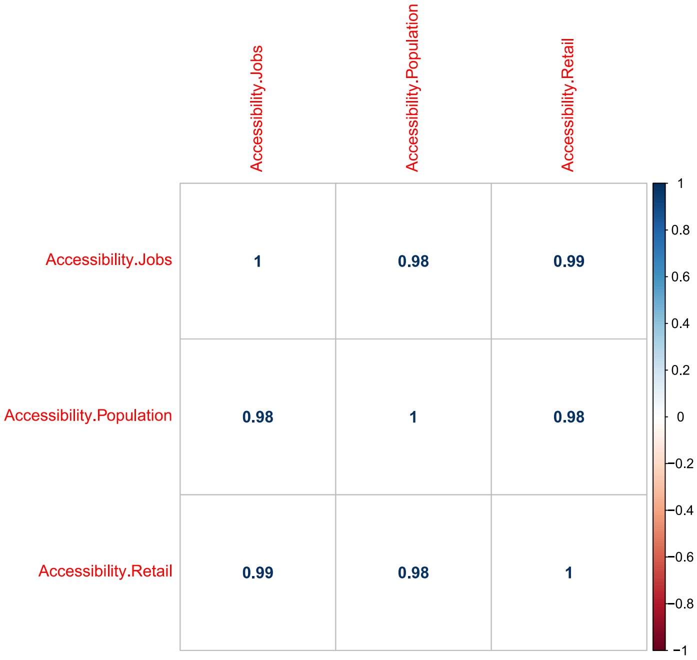

Fig. 21: correlation of accessibility to different destinations

4. Model estimation and results

For bid function estimation, households were aggregated based on their incomes. Each of the aggregated households share the same bid function coefficient values. Initially, there were 18 different income groups in the study area. Firstly, they were aggregated into six income groups and then three income groups. The model of three income groups outperformed six income groups, so the estimation was proceeded for three income groups (Appendix E). The utility of these three households can be written as below:

Eq. 4-1:

$$\begin{aligned} B_h = \beta_0 + & \sum \beta_{Zonal\ Attributes} * Zonal\ Attribues + \\ & \sum \beta_{Household\ Attributes} * Household\ Attribues + \\ & \sum \beta_{Real\ estate\ Attributes} * Real\ estate\ Attribues \end{aligned}$$

Eq. 4-2:

$$\begin{aligned} B_{income\_1} = \beta_{0\_1} + & \beta_{1\ Accessibility} * Accessibility + \\ & \beta_{1\ Urban\ or\ Rural} * Urban\ or\ Rural + \\ & \beta_{1\ Number\ of\ schools} * Number\ of\ schools^{1} + \\ & \beta_{1\ \%\ High\ income * High\ income} * \%\ High\ income * High\ income + \\ & \beta_{1\ \%\ Low\ income * High\ income} * \%\ Low\ income * High\ income + \\ & \beta_{1\%\ Educated * Education} * \%\ Educated * Education + \\ & \beta_{1\ \%\ Race * Race} * \%\ Race * Race^{2} + \\ & \beta_{1\ Income\ zonal} * Income\ zonal + \\ & \beta_{1 Retired} * Retired + \\ & \beta_{1\ Gender} * Gender + \\ & \beta_{1\ Home\ type * household\ size} * Home\ type * household\ size \end{aligned}$$

$$\begin{aligned} B_{incom\_2} = \beta_{0\_2} + & \beta_{2\ Accessibility} * Accessibility + \\ & \beta_{2\ Urban\ or\ Rural} * Urban\ or\ Rural + \\ & \beta_{2\ Number\ of\ schools} * Number\ of\ schools + \\ & \beta_{2\ \%\ High\ income * High\ income} * \%\ High\ income * High\ income + \\ & \beta_{2\ \%\ Low\ income * High\ income} * \%\ Low\ income * High\ income + \end{aligned}$$

[1] Or independent variables: Number of schools, or Children, these parameters have high correlation, so cannot enter to the model at the same time
[2] Or race

$$\beta_{2\%\,Educated\,*\,Education} * \%\,Educated * Education +$$
$$\beta_{2\,\%\,Race\,*\,Race} * \%\,Race * Race +$$
$$\beta_{2\,Income\,zonal} * Income\ zonal +$$
$$\beta_{2Retired} * Retired +$$
$$\beta_{2\,Gender} * Gender +$$
$$\beta_{2\,Home\,type\,*\,household\,size} * Home\ type * household\ size$$

$$B_{incom\_3} = \beta_{0\_3} + \beta_{3\,Accessibility} * Accessibility +$$
$$\beta_{3\,Urban\,or\,Rural} * Urban\ or\ Rural +$$
$$\beta_{3\,Number\,of\,schools} * Number\ of\ schools +$$
$$\beta_{3\,\%\,High\,income\,*\,High\,income} * \%\,High\ income * High\ income +$$
$$\beta_{3\,\%\,Low\,income\,*\,High\,income} * \%\,Low\ income * High\ income +$$
$$\beta_{3\%\,Educated\,*\,Education} * \%\,Educated * Education +$$
$$\beta_{3\,\%\,Race\,*\,Race} * \%\,Race * Race +$$
$$\beta_{3\,Income\,zonal} * Income\ zonal +$$
$$\beta_{3Retired} * Retired +$$
$$\beta_{3\,Gender} * Gender +$$
$$\beta_{3\,Home\,type\,*\,household\,size} * Home\ type * household\ size$$

Logit model does not provide absolute estimations of the utilities. This model is able to give relative values for willingness to pay, which is relative to one of the utilities (Hurtubia and Bierlaire, 2013). Normally, the alternative with the highest frequency will be normalized to zero. Based on Fig. 22, income group 3 has the highest frequency, so the constant of this alternative was set to be zero and constant of other alternatives will be estimated relative to this alternative.

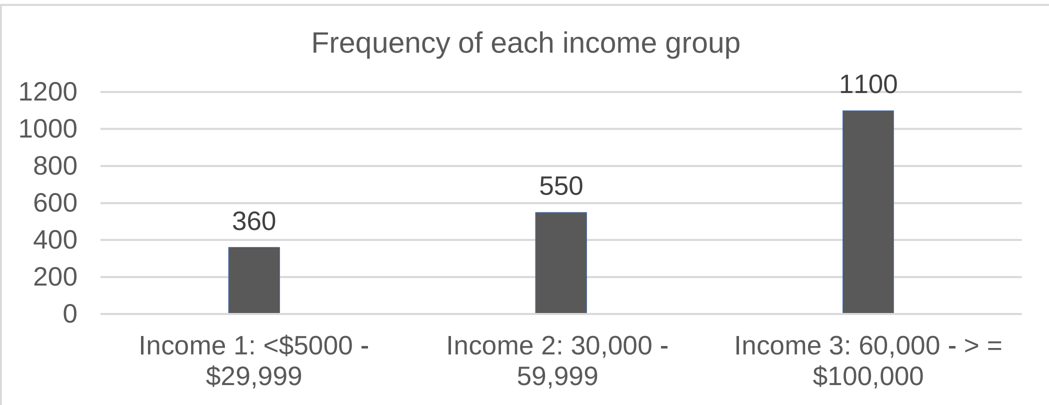

Fig. 22: Frequency of each income group in NHTS

The measurements will be done based on Eq. 4-10. The total number of each income group) was calculated from the census data of 2014 to be assumed as H_h (Table 8).

$$P_{h/vi} = \frac{exp(\mu\ B_{hvi} + ln(H_h))}{\sum_g exp(\mu\ B_{vi} + ln(H_g))} \qquad \text{Eq. 4-10}$$

Table 8: Frequency of each income group in census data

Household Group	Frequency	Ln(Frequency)
Income_1	106,857	11.57
Income_2	126,873	11.75
Income_3	219,600	12.29

The estimation was repeated for five different accessibility measurements, three different gravity accessibility measurements and two different cumulative accessibility measurements. Calculation of accessibilities was for three different destinations, jobs, population and retail jobs, so the estimation was repeated 15 times.

To estimate β parameters of the model, the Biogeme that uses maximum likelihood estimation method to estimate parameters of Logit model was used (Bierlaire, 2003). The NHTS's reported weights were used during the estimation to adjust the survey and to produce reliable estimates (NHTS, 2009).

4.1 Results and discussion

The estimation was done for all the explanatory variables based on Eq. 5-2. Then, the variables that were not statistically significant[1] dropped from the estimation. Following, there is an explanation about the findings during estimations:

- **Education:** This variable was showing negative value that means if the households are educated, they are less likely to be selected by landlord. This result does not seem explainable, so this variable removed from the final estimation.
- **Race:** this variable due to negative value dropped from the estimation. The negative value means households do not value zones where their race is dominant. This result was not explainable either.

[1] Statistical significance at the 10% level means p-value is less than 0.1

- **Accessibility:** this parameter was assumed to be significant for both group of households. Some accessibility measurements were important for both income groups, but some were important for one of the income groups.
- **% high income * high income, % Low income * High income, % Educated * Education, % Race * Race, Zonal income, Number of schools, Number of schools * Child:** although at the beginning it was assumed that these parameters will be significant, but they dropped during the estimation due to their insignificancy.

The estimation was repeated for 15 different accessibility measurements. These 15 different outputs of the estimations were compared with each other based on their final log-likelihood (Table 8). It is assumed that the one with the smallest final log-likelihood, closer to zero, is the best representor of the study area (Hurtubia and Bierlaire, 2013). Table 8 shows the explanatory variables of the model. The results presented in Table 8 belong to traditional cumulative accessibility measurement for retail jobs (Eq. 3-12). Later, it will be shown that this accessibility measurement outperforms other accessibility measurements. The statistical summary of the estimation is given in Table 9.

Table 8: Estimation results for Richmond, Virginia

Income group	Variable	Value	t-test
Income_1	Constant_1	-0.107	-0.78
Income_2	Constant_2	-0.0207	-0.12
Income_3	Constant_3	0.0[1]	
Income_1	Accessibility_1[2]	3.35	7.94
Income_2	Accessibility_2	1.04	2.44
Income_1	Children_1	-0.329	-2.01
Income_2	Children_2	-0.617	-4.66
Income_1	Gender_1	-1.19	-9.61
Income_2	Gender_2	-0.474	-4.30
Income_2	Home type * household size_2	0.0627	2.04
Income_1	Retired_1	1.43	9.64
Income_2	Retired_2	0.438	3.34
Income_2	Urban or Rural_2	0.317	2.83

[1] Under-identification of Logit model, the constant for income level 3 is manually fixed to zero (Hurtubia and Bierlaire, 2013)
[2] Destination Summation accessibility measurement to retail jobs

Table 9: Statistical summary of the estimation

Parameter	Estimate
Model	Logit
Number of observations	2010
Final log likelihood	-1941.270
Rho-square	0.121

Following, there are explanations about the significant variables:

Accessibility: As it was predictable, accessibility is one of the explanatory variables for both income groups. The positive value of this parameter indicates that zones with high accessibility to retail jobs attract both groups of households. Importance of accessibility to retail jobs was observed also in the other studies for Boston and Eugene-Springfield (Waddell, 2000; Citilabs, 2013).

Children: Households with children in both income groups are less likely to win real estate auctions. This means that landlords are not willing to choose households with children.

Gender: The gender parameter takes 1 for male gender, so the results are related to the male households. Gender is an explanatory variable for both income groups. The negative sign of the gender indicates that in households that the head of household is male have less chance to win the real estate auctions.

Home type * household size_2: This dummy parameter tests whether larger households prefer to live in detached dwellings or not. The result indicates that larger household in income group 2 desire to live in detached dwellings. This result was observed also in a study in Brussels (Hurtubia and Bierlaire, 2013).

Retired: This parament indicates that how landlords react to retired households and is significant for both income groups. The positive value of it reveals that retired households have more chance to win the real estate auctions.

Urban or Rural_2: This parameter is significant for income group 2. The positive sign of it shows that households of income group 2, richer households, prefer to locate in urban areas.

Constants: Constant of the utility function is the expected mean that the utility function gets when all independent variables are equal to zero and it is a representor for the variables that were not included in the utility function. The constants of this model do not appear to be significant, it can be assumed that the variables that were considered in the model do not explain the choice completely. However, the negative sign for constants could be due to the positive contribution of existing variable to the utility of the alternatives.

4.2 Accessibility measurements' comparison

As stated earlier, accessibility measurements for this study include three different gravity accessibility measurements and two different cumulative accessibility measurements for three different destinations: jobs, population and retail jobs. A brief comparison between these accessibility measurements and their significance is given in Table 10. Gravity accessibility measurement based on Eq. 2-2 for jobs, population and retail jobs, plus, gravity accessibility measurement based on Eq. 2-3 for jobs and for retail jobs, plus, gravity accessibility Eq. 2-4 for jobs and retail jobs were not significant. In other words, their t-tests were not great, and they dropped during the estimations.

All added accessibility measurements in Table 10 with their final log-likelihood were significant. The final log-likelihood and R^2 of these significant accessibility measurements differ slightly from each other. However, the traditional cumulative accessibility measurement for retail jobs has the best final log-likelihood[1], which indicates that this accessibility is the best fit to the estimation data (Hurtubia and Bierlaire, 2013). This cumulative accessibility measurement outperforms other accessibility measurements, which is immediately followed by gravity Eq. 2-4 accessibility measurement for population. Furthermore, just these two accessibility measurements were significant for both income groups and all the other were significant for the first income group.

The power of cumulative accessibility measurement in explaining residentials' location choice was also showed by Srour et al. (Srour et al., 2002). This finding is helpful in defining transportation policies. For example, in a location if accessibility for retail jobs, increases in a way that they are accessible within 20 minutes by car[2], more households will attract to locate in that location.

[1] Closer to zero
[2] Refer to assumptions in calculating traditional cumulative accessibility measurement

Access score or weighted cumulative accessibility measurement is significant for all three destinations and for income group one. It shows that low income households (income_1) not only prefer good accessibility to jobs, population and retail jobs but also in their perception of accessibility, they value closer opportunities more.

Table 10: Estimation results for different accessibility measurements

Accessibility		Final log-likelihood	R^2	Significance	Income groups
Gravity					
	Jobs	-	-	Not significant	
Based on Eq. 2-2	Population	-	-	Not significant	
	Retail jobs	-	-	Not significant	
	Jobs	-	-	Not significant	
Based on Eq. 2-3	Population	-1978.919	0.104	-	Income_1
	Retail jobs	-	-	Not significant	
	Jobs	-	-	Not significant	
Based on Eq. 2-4	**Population**	**-1963.762**	**0.111**	-	**Income_1 & _2**
	Retail jobs	-	-	Not significant	
Destination Summation	Jobs	-1973.569	0.106	-	Income_1
	Population	-1972.639	0.107	-	Income_1
	Retail jobs	**-1941.27**	**0.121**	-	**Income_1 & _2**
	Jobs	-1977.246	0.105	-	Income_1
Access Score	Population	-1977.07	0.105	-	Income_1
	Retail jobs	-1981.268	0.103	-	Income_1

4.3 Gravity Eq. 2-4 accessibility measurement

As it is stated earlier, gravity Eq. 2-4 accessibility measurement for population is one of the significant accessibility measurements. This accessibility measurement was calculated for different α and β. α defines the weight of populated areas and β is a parameter with negative sign to reflect that the nearby destinations get greater weight. The result of estimations revealed that the final log-likelihood of these four gravity Eq. 2-4 accessibilities differ slightly from each other (Table 11), however, the final log-likelihood of the first model with $\alpha = 1.0$ and $\beta = 0.5$ which means high impact of travel time and low impact of urban centers has the best fit to the estimation data. In other words, accessibility measurement which was calculated based on assuming that households value more in travel time to populated zones rather than amount of the population in the zone describes better households' interpretation of accessibility. These accessibility measurements were visualized in Fig. 23, Fig. 24, Fig. 25 and Fig. 26.

Table 11: Estimation results for gravity Eq. 2-4 accessibility measurement to population

α	**β**	**Final log-likelihood**	**R^2**	**impact of the two parameters α and β**
1.0	**0.5**	**-1963.762**	**0.111**	High impact of travel time, low impact of urban centers
1.0	0.3	-1971.342	0.107	Low impact of travel time, low impact of urban centers
1.5	0.5	-1977.364	0.105	High impact of travel time, high impact of urban centers
1.5	0.3	-1981.204	0.103	Low impact of travel time, high impact of urban centers

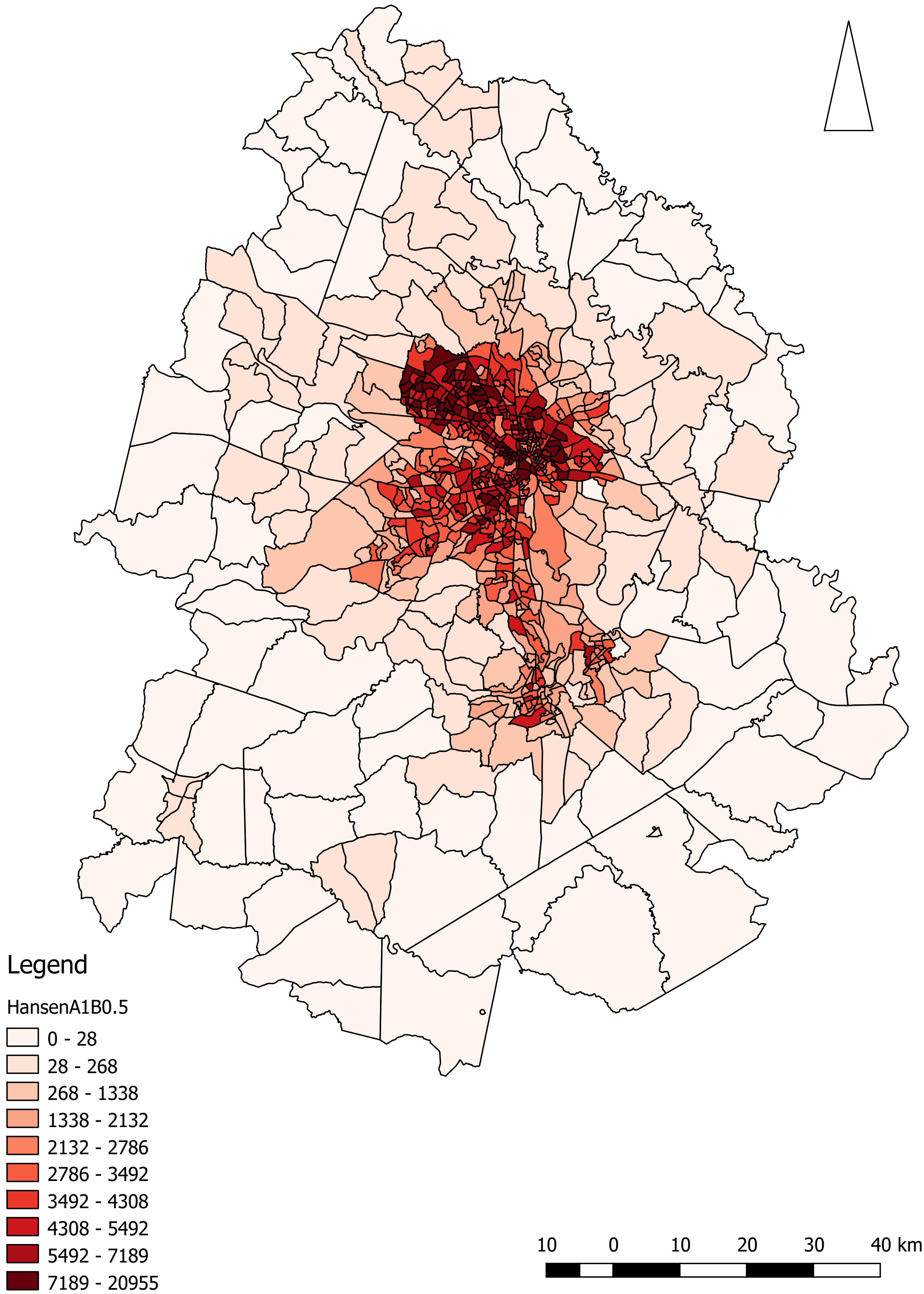

Fig. 23: Gravity Eq. 2-4 accessibility measurement with $\boldsymbol{\alpha} = \mathbf{1.0}, \boldsymbol{\beta} = \mathbf{0.5}$

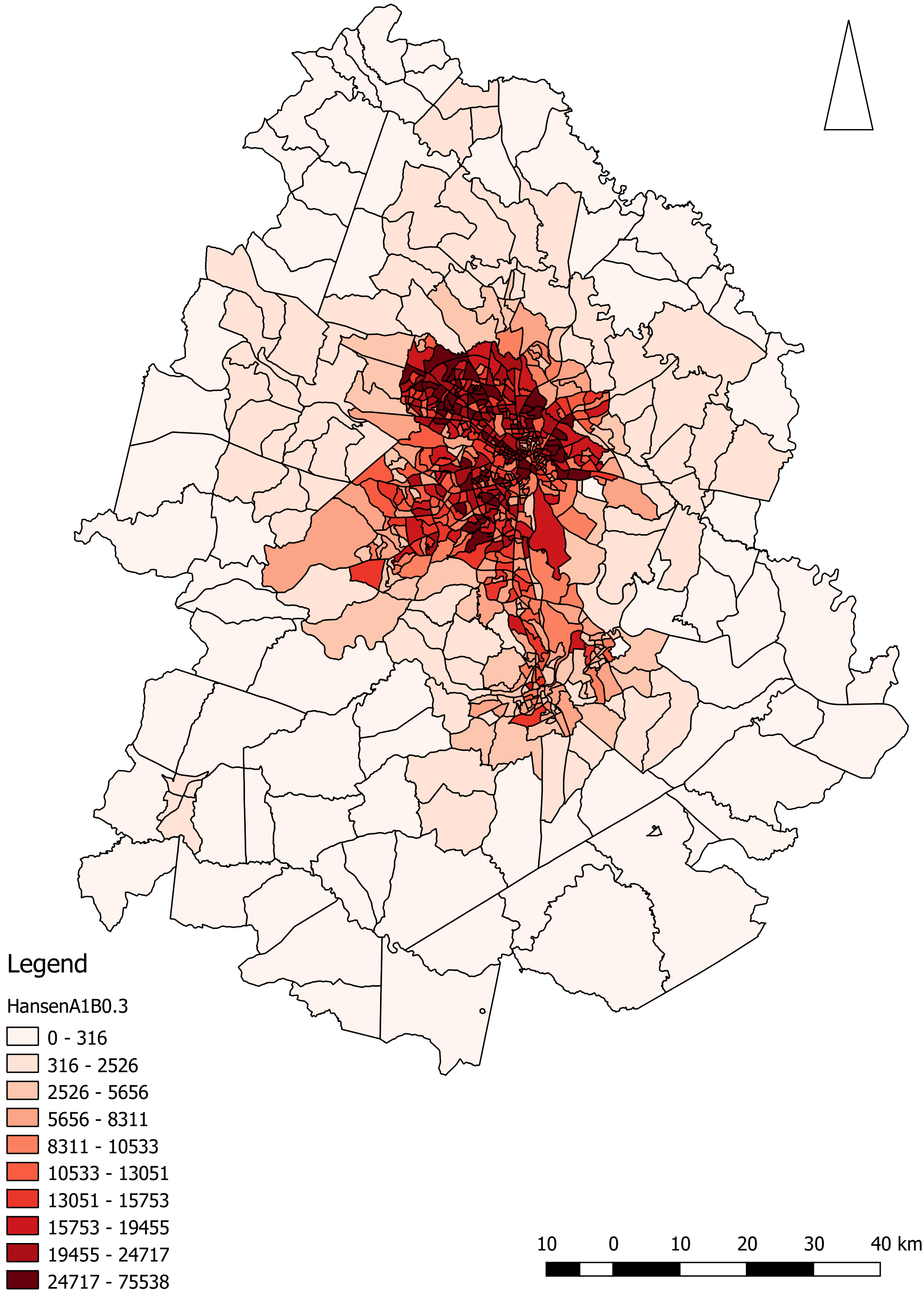

Fig. 24: Gravity Eq. 2-4 accessibility measurement with $\alpha = 1.0, \beta = 0.3$

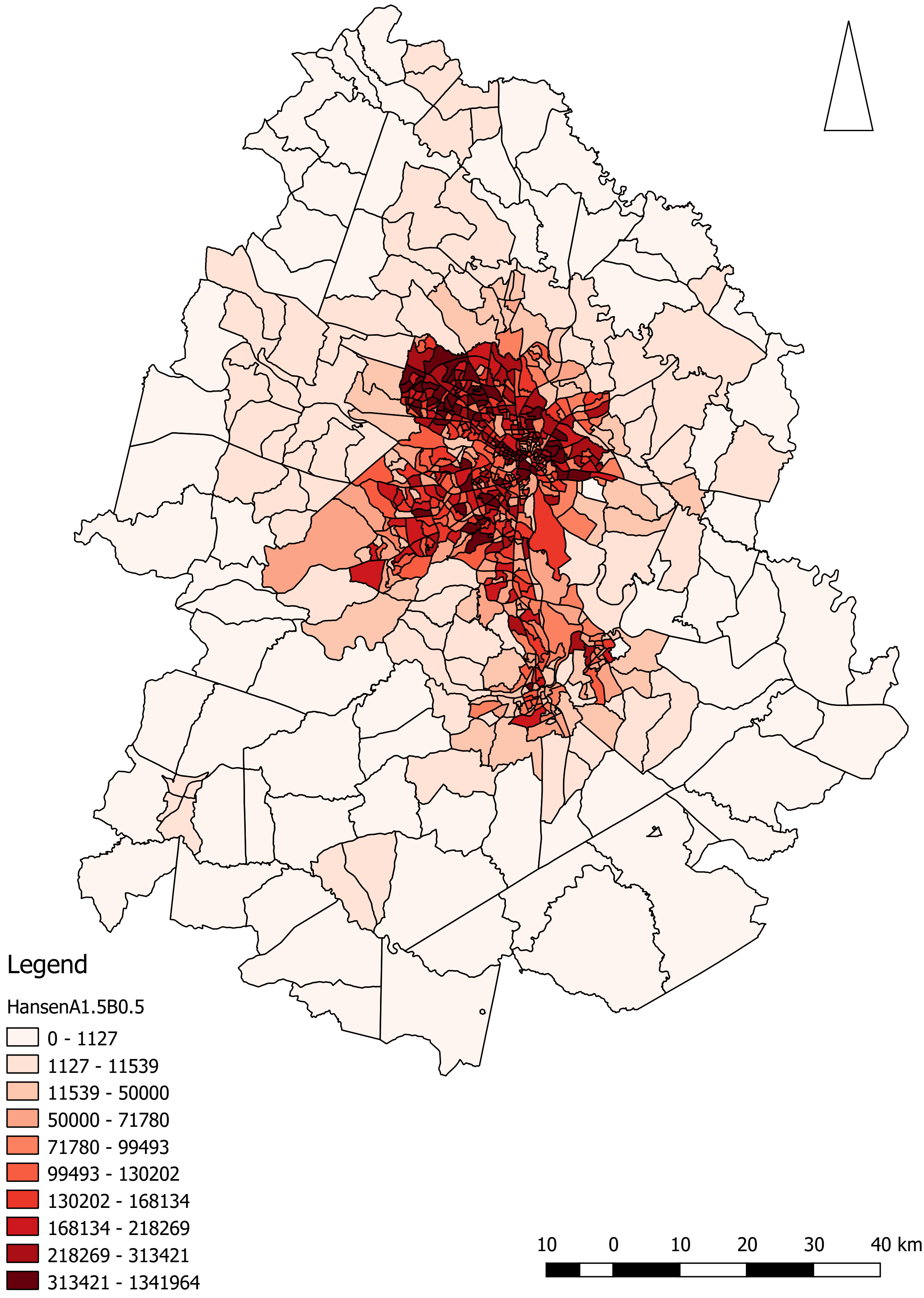

Fig. 25: Gravity Eq. 2-4 accessibility measurement with $\alpha = 1.5, \beta = 0.5$

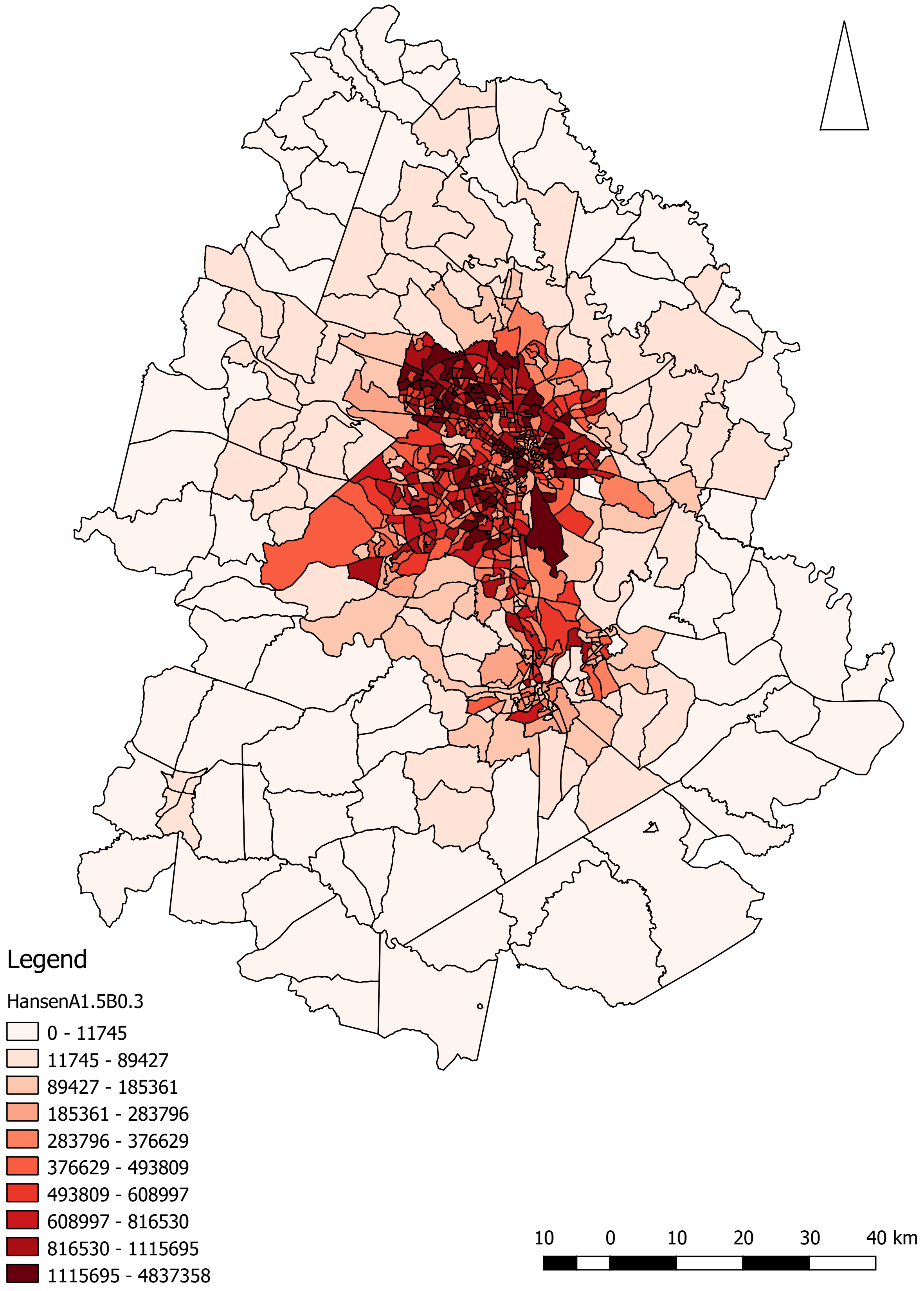

Fig. 26: Gravity Eq. 2-4 accessibility measurement with $\alpha = 1.5, \beta = 0.3$

5. Conclusion

The aim of this thesis was to assess the role of accessibility in households' location choice. To do this, the census 2010 and NHTS 2009 were used to determine the zonal attributes, the household attributes and the real estate attributes of each household surveyed in the study area, Richmond, Virginia. The altitude and latitude of the households surveyed by NHTS were used to place each household in the ArcMap to determine the zonal attributes. The household attributes and the real estate attributes were extracted directly from the NHTS. To measure accessibility, which was assumed to be a zonal attribute, the Sugar Access of Citilabs, an application for ArcGIS, was used. Accessibility to three different destinations-- jobs, population and retail jobs-- and five different accessibility measurement methods were chosen to be tested for this thesis. These methods include three different gravity accessibility measurements and two different cumulative accessibility measurements including the weighted cumulative accessibility measurement, meaning that nearer destinations were weighted more than farther destinations. Overall, 15 accessibility measurements were taken for this thesis.

A bid-auction approach to modeling households' location choice was estimated with the help of an aggregated Logit method (Martinez, 1992). This Logit model was estimated with the Biogeme package (Bierlaire, 2003) . In this method, aggregated households are alternatives, so the households were aggregated based on their incomes. It is assumed that these aggregated households share bid function coefficient values.

In contrast to earlier predictions, the importance of many parameters dropped due to their insignificance. Examples of these would be the race and educational status of households, average zonal income, the percentage of educated people and the percentage of people with the same race as the household in the zone. However, after estimating different attributes, the following attributes were used to determine the model: accessibility, existence of children, household's gender, retirement status, house type and urbanization of the zone. Households generally prefer to locate in urbanized zones with more accessibility while larger households are more interested in living in detached dwellings. Furthermore, retired households have higher chances of winning the real estate auctions. Landlords are not willing to choose households with children and households with a male head.

In the next step of this thesis, different accessibility measurements methods were examined. Some of the methods, such as gravity based on Eq. 2-2, accessibility measurements of all the destinations mentioned, gravity based on Eq. 2-3, and Eq. 2-4 accessibility measurements for jobs and retail jobs, were not significant. Cumulative accessibility measurements[1] and weighted cumulative accessibility measurements[2] for all destinations were significant. Furthermore, gravity based on Eq. 2-3 and Eq. 2-4 accessibility measurements for population were also statistically significant. The final log-likelihoods of these significant accessibility measurements were slightly different. However, the final log-likelihood of the cumulative accessibility measurement for retail jobs outperforms other accessibility measurements, which is immediately followed by the gravity Eq. 2-4 accessibility measurement for population. It means in a location defining new policies in a way to increase accumulative accessibility to retail jobs will lead to higher attraction of the location for the households.

In another estimate, gravity Eq. 2-4 accessibility measurements for population with four different calibration parameters were modeled. These calibration parameters clarify the impacts of travel time and urban centers. The result of the estimations shows that all these four models are statistically significant. Though, the gravity Eq. 2-4 accessibility with $\alpha = 1.0$ and $\beta = 0.5$ outperforms other gravity Eq. 2-4 accessibility measurements, which means high impact of travel time and low impact of urban centers explain the best fit to the estimation data.

5.1 Limitations and suggestions

The main limitation of this study was related to the available data and number of recorded households in NTHS data. Some of the limitations and suggestions of this study can be listed as below:

- Three groups of attributes were needed to model households' location choice: zonal attributes, household attributes and real estate attributes. The real estate attributes such as dwelling size, age, number of bathrooms, number of bedrooms, number of rooms were missing in dataset. Dwelling type, was the only real estate attribute that was available for the study area of this thesis. The estimations revealed the importance of dwelling type. Accordingly, other real estate attributes might also be explanatory.
- Increasing the sample size by collecting more recorded households would lead to better estimation of households' location choice (Bierlaire, 2016). Only 2010 surveyed household were available in the study area. Furthermore, the two available sources for the study area of this thesis were from different years. The census data was from 2010 and the NTHS

[1] Called Destination Summation by Citilabs
[2] Called Access Score by Citilabs

was from 2009. It was assumed that this difference does not affect the purpose of this thesis.

- The study area, Richmond, Virginia, was divided into 746 zones. The NTHS was covering households from 605 zones and there was no available household surveyed in the rest of 141 zones. Improving the data by covering all zones of the study area will improve the output.
- For this study, households were aggregated based on their income. Applying other methods of aggregating households can help in better understanding of households' location choice.
- In the case of availability, other parameter such as number of crime, air quality and accessibility to recreational, shopping centers can be checked.
- Validation of the model was not within the scope of this thesis. However, it is necessary for further use of outcomes of this work.

References

Alonso, W., 1964. Location and Land Use. Harvard University Press.

Ben-Akiva, M., Bowman, J.L., 1998. INTEGRATION OF AN ACTIVITY-BASED MODEL SYSTEM AND A RESIDENTIAL LOCATION MODEL. SAGE J.

Ben-Akiva, M., Lerman, S.., 1977. Disaggregate travel and mobility choice models and measures of accessibility. Presented at the Proceedings of the Third International Conference on Behavioral Travel Modeling, Tanenda, Australia.

Bierlaire, M., 2016. Common mistakes in discrete choice modeling. Episode I: To be or not to be significant.

Bierlaire, M., 2003. BIOGEME: a free package for the estimation of discrete choice models. Presented at the Conference paper STRC, EPFL, p. 27.

Bina, M., Kockelman, K., Suescun, D., 2006. Location choice vis-á-vis transportation: the case of recent home buyers, in: Proceedings of the 85th Annual Meeting of the Transportation Research Board.

Brian H., Y.L., 2009. Re-examining the influence of work and non-work accessibility on residential location choices with a micro-analytic framework.

Cascetta, E., 2001. Transportation Systems Engineering: Theory and Methods. Springer.

Census, 2010. What is the Census? [WWW Document]. URL https://www.census.gov/2010census/about/

Chattopadhyay, S., 1997. An Empirical Investigation into the Performance of Ellickson's Random Bidding Model, with an Application to Air Quality Valuation. J. Urban Econ. 43, 292–314. https://doi.org/10.1006/juec.1997.2046

Cirilabs, 2016. SUGAR ACCESS: METHODOLOGY AND APPLICATION. Citilabs Inc.

Citilabs, 2016. Sugar Access User Guide. Citilabs Inc.

Citilabs, 2014. Cube Land User Guide.

Citilabs, 2013. Cube Land Boston User Guide.

Dong, H., Gliebe, J., 2011. Forecasting the Location of New Housing in Integrated Land Use 2 Models: Comparison of Three Approaches to the Developer's 3 Perspective in the Portland Region. Transp. Res. Board.

Dong, X., Ben-Akiva, M.E., Bowman, J.L., Walker, J.L., 2006. Moving from trip-based to activity-based measures of accessibility. Transp. Res. Part Policy Pract. 40, 163–180. https://doi.org/10.1016/j.tra.2005.05.002

El-Geneidy, A.M., Levinson, D.M., 2006a. Access to destinations: Development of accessibility measures.

El-Geneidy, A.M., Levinson, D.M., 2006b. Access to Destinations Study (No. 1), 2006-16.

Ellickson, B., 1981. An alternative test of the hedonic theory of housing markets. J. Urban Econ. 9, 56–79.

Esri, n.d. ArcGIS.

Geurs, K., 2006. Accessibility, land use and transport. EBURON.

Geurs, K.T., van Wee, B., 2004. Accessibility evaluation of land-use and transport strategies: review and research directions. J. Transp. Geogr. 12, 127–140. https://doi.org/10.1016/j.jtrangeo.2003.10.005

Glickman, I., Katoshevski, R., Ishaq, R., Shiftan, Y., 2015a. Integrating activity-based travel-demand models with land-use and other long-term lifestyle decisions. J. Transp. Land Use. https://doi.org/10.5198/jtlu.2015.658

Glickman, I., Katoshevski, R., Ishaq, R., Shiftan, Y., 2015b. Integrating activity-based travel-demand models with land-use and other long-term lifestyle decisions. J. Transp. Land Use. https://doi.org/10.5198/jtlu.2015.658

Guo, J., Bhat, C., 2001. Residential location choice modeling: Accommodating sociodemographic, school quality and accessibility effects. Univ. Tex. Austin.

Guo, J.Y., Bhat, C.R., 2007. Operationalizing the concept of neighborhood: Application to residential location choice analysis. J. Transp. Geogr. 15, 31–45.

Horner, M.W., 2004. Exploring Metropolitan Accessibility and Urban Structure. Urban Geogr. 25, 264–284. https://doi.org/10.2747/0272-3638.25.3.264

Hurtubia, R., Bierlaire, M., 2013. Estimation of Bid Functions for Location Choice and Price Modeling with a Latent Variable Approach. Netw. Spat. Econ. 14, 47–65. https://doi.org/10.1007/s11067-013-9200-z

Hurtubia, R., Martinez, F., Flötteröd, G., Bierlaire, M., 2010. Comparative analysis of hedonic rents and maximum bids in a land-use simulation context, in: 10th Swiss Transport Research Conference.

Hwang, S., Thill, J.-C., 2010. Influence of Job Accessibility on Housing Market Processes: Study of Spatial Stationarity in the Buffalo and Seattle Metropolitan Areas, in: Jiang, B., Yao, X. (Eds.), Geospatial Analysis and Modelling of Urban Structure and Dynamics. Springer Netherlands, Dordrecht, pp. 373–391. https://doi.org/10.1007/978-90-481-8572-6_19

Jara-Díaz, S.R., Martinez, F.J., 1999. On the specification of indirect utility and willingness to pay for discrete residential location models. J. Reg. Sci. 39, 675–688.

Karst, T., van Eck, J.R.R., 2003. Evaluation of Accessibility Impacts of Land-Use Scenarios: The Implications of Job Competition, Land-Use, and Infrastructure Developments for the Netherlands. Environ. Plan. B Plan. Des. 30, 69–87. https://doi.org/10.1068/b12940

Kockelman, K.M., Kalmanje, S., 2003. CREDIT-BASED CONGESTION PRICING: TRAVEL, LAND VALUE AND WELFARE IMPACTS 6. Performing Organization Code.

Lee, B.H., Waddell, P., Wang, L., Pendyala, R.M., 2009. Operationalizing time-space prism accessibility in a building-level residential choice model: empirical results from the Puget Sound region, in: Transportation Research Board 88th Annual Meeting.

Lee, B.H.Y., Waddell, P., 2010. Residential mobility and location choice: a nested logit model with sampling of alternatives. Transportation 37, 587–601. https://doi.org/10.1007/s11116-010-9270-4

Lee, B.H.Y., Waddell, P., Wang, L., Pendyala, R.M., 2010. Reexamining the Influence of Work and Nonwork Accessibility on Residential Location Choices with a Microanalytic Framework. Environ. Plan. A 42, 913–930. https://doi.org/10.1068/a4291

Lerman, S.R., Kern, C.R., 1983. Hedonic theory, bid rents, and willingness-to-pay: some extensions of Ellickson's results. J. Urban Econ. 13, 358–363.

Martinez C., F.J., 1995. ACCESS: THE TRANSPORT-LAND USE ECONOMIC LINK. ScienceDirect.

Martínez, F., 1996. MUSSA: Land Use Model for Santiago City. Transp. Res. Board.

Martínez, F., Araya, C., 2000. Transport and Land-Use Benefits under Location Externalities. Environ. Plan. A 32, 1611–1624. https://doi.org/10.1068/a32131

Martinez, F.J., 1992. The bid—choice land-use model: An integrated economic framework. Environ. Plan. A 24, 871–885.

McFadden, D., 1978. Modeling the Choice of Residential Location. Transp. Res. Board.

Miyamoto, A., Kitazume, K., 1989. land-use model based on random utility/rent-bidding analysis (rurban). Presented at the Transport policy management and technology - towards 2001 selected proceedings of the fifth world conference on transport research, Yokohama.

Moeckel, R., 2017. Accessibilities [WWW Document]. Travel Forecast. Resour. URL http://tfresource.org/Accessibilities

NHTS, 2009.

Nuzzolo, A., Coppola, P., 2007. Accessibility and socioeconomic activities location, in: Proceedings of European Transportation Conference.

R, Package ‘corrplot,’ 2017.

Salvini, P., Miller, E.J., 2005. ILUTE: An Operational Prototype of a Comprehensive Microsimulation Model of Urban Systems. Netw. Spat. Econ. 5, 217–234. https://doi.org/10.1007/s11067-005-2630-5

Schürmann, C., Spiekermann, K., Wegener, M., 1997. Berichte aus dem Institut für Raumplanung 39, Accessibility Indicators.

Shen, Q., 1998. Location characteristics of inner-city neighborhoods and employment accessibility of low-wage workers. Environ. Plan. B Plan. Des. 25, 345–365. https://doi.org/10.1068/b250345

Srour, I.M., Kockelman, K.M., Dunn, T.P., 2002. Accessibility Indices: A Connection to Residential Land Prices and Location Choices. Transp. Res. Rec. J. Transp. Res. Board.

Wachs, M., Kumagai, T.G., 1973. PHYSICAL ACCESSIBILITY AS A SOCIAL INDICATOR. ScienceDirect.

Waddell, P., 2000. A Behavioral Simulation Model for Metropolitan Policy Analysis and Planning: Residential Location and Housing Market Components of Urbansim. Environ. Plan. B Plan. Des. 27, 247–263. https://doi.org/10.1068/b2627

Weber, J., 2003. Individual accessibility and distance from major employment centers: An examination using space-time measures. J. Geogr. Syst. 5, 51–70.

Wegener, M., Fürst, F., 1999. Land-use transport interaction: state of the art.

Williams, I., 1976. A comparison of some calibration techniques for doubly constrained models with an exponential cost function. Transp. Res. 10, 91–104.

Wilson, A.G., 1970. Entropy in urban and regional modelling. Routledge Revivals.

Wilson, A.G., 1967. A statistical theory of spatial distribution models. Transp. Res.

Wu, W., Zhang, W., Dong, G., 2013. Determinant of residential location choice in a transitional housing market: Evidence based on micro survey from Beijing. Habitat Int. 39, 16–24. https://doi.org/10.1016/j.habitatint.2012.10.008

Zondag, B., Pieters, M., 2005. Influence of accessibility on residential location choice. Transp. Res. Rec. J. Transp. Res. Board 63–70.

Appendixes

Appendix A: Census data

Field	Definition	Description
OBJECTID	ID number	ID number for each link from the original shapefile.
STATEFP10	State ID	Census State ID code
COUNTYFP10	County ID	Census County ID code
TRACTCE10	Census Tract ID	Census Tract ID code
BLOCKCE	Census Block ID	Census Block ID code
BLOCKID	Census Block Whole ID	Unique Census Block ID code
HOUSEHOLDS	Number of Households	Number of households in Census Block
HHR##-##	Age of Household Owner	Number of heads of household from age ## to age ##
HH#	Household Population	Number of households with # people
POPULATION	Total Population	Total population
WHITE	White Population	Total White population
BLACK	Black Population	Total Black population
AMERICAN_INDIAN	American Indian Population	Total American Indian population
ASIAN	Asian Population	Total Asian Population
PACIFIC_ISLAND	Pacific Island Population	Total Pacific Island population
OTHER	Other Population	Total population of other races
A#_TO_#	Age of Population	Population of Census Block from age ## to age ##
SCHOOLKIDS	School Aged Kids	Number of kids age 5-17
SENIORS	Senior Aged People	Number of people age 65 and up
JOBS	Number of Jobs	Total number of jobs
EMP_TO29	Workers Under 29	Number of jobs for worked age 29 or younger
EMP_30_54	Workers 30-54	Number of jobs for workers age 30 to 54
LOWINCJOBS	Low Income Jobs	Number of jobs with earnings $1250/month or less
MEDINCJOBS	Medium Income Jobs	Number of jobs with earnings $1251/month to $3333/month
HIGHINCJOBS	High Income Jobs	Number of jobs with earnings greater than $3333

Appendix B: Dataset Example

ID ZONE	Weight	Income	Accessibility	Urban or Rural	Nr.Schools	NrSchool*Child	IncomeZo	Retired	Children	Age	Gender	Education	Race	HomeType*HI	%HighIncome*Hi
249	0.88	18	60	0	0	0	1	0	1	39	1	1	0	2	100
18	0.61	18	29	1	0	0	0	0	0	59	1	1	1	0	80.21
159	0.34	17	59	1	1	0	1	1	0	54	1	0	1	0	88.49
89	0.53	12	73	0	0	0	0	1	0	76	2	0	0	1	74.6
589	0.88	12	99	1	0	0	0	0	0	56	2	1	1	0	82.04
390	1.02	10	86	1	0	0	0	0	0	40	2	0	1	0	68.34
246	0.54	18	49	1	0	0	1	0	1	52	1	0	0	0	90.49
715	0.45	17	86	1	1	1	1	0	1	54	1	1	0	4	90.91
47	1.15	1	61	0	0	0	0	0	0	63	1	0	1	1	0
556	1.35	6	34	1	1	0	0	1	0	79	1	0	1	0	0
488	0.64	17	23	1	1	1	0	0	1	45	2	0	0	0	80
433	0.25	9	53	1	0	0	0	1	0	61	1	0	0	0	93.44
370	0.22	13	73	1	2	0	1	1	0	74	1	1	1	2	89.68
248	0.44	18	60	0	1	0	1	0	0	76	2	0	0	2	86.35
656	1.05	12	99	0	0	0	0	0	1	24	1	0	0	2	72.18
251	0.61	18	68	1	0	0	1	0	0	45	1	0	1	0	97.12
73	0.49	18	53	1	1	1	1	0	1	48	1	1	0	1	96.54
657	1.76	18	92	1	0	0	1	0	1	42	2	1	0	2	80.09
265	1.00	5	70	1	0	0	0	1	0	90	1	0	0	0	0
86	0.30	14	29	1	1	1	1	0	1	47	1	1	0	0	95.94
709	0.70	11	79	1	4	0	1	0	0	41	1	1	1	1	87.85

Appendix C: Example of explanatory variables used in similar studies

Title	Boston Region Land Use Model (Citilabs, 2013)	A behavioral simulation model for metropolitan policy analysis and planning: residential location and housing market components of UrbanSim (Waddell, 2000)	Estimation of Bid Functions for Location Choice and Price Modeling with a Latent Variable Approach (Hurtubia and Bierlaire, 2013)	An Empirical Investigation into the Performance of Ellickson's Random Bidding Model, with an Application to Air Quality Valuation (Chattopadhyay, 1997)
Publish year	2013	2000	2013	1997
Researchers	Citilabs Inc.	Paul Waddell	Ricardo Hurtubia and Michel Bierlaire	Sudip Chattopadhyay
Approach	Bid-auction	Bid-auction	Bid-auction	Bid-auction
Estimation method	Martinez	Martinez	Ellickson	Elickson
Place	Boston/USA	Lane County/USA	Brussels/Belgium	Chicago/USA
Explanatory variables	• household size • age of head of household • a dummy variable for low household income • drive access to retail jobs • of drive access to service jobs • number of rooms per unit • multi-family indicator • transit access • drive access to basic sector employment	• housing type • accessibility to total employment • accessibility to retail employment • density of a particular housing type in a zone • number of housing units of a particular type in the zone • average age of the buildings • Percentage of households in a zone in different income group • Travel time to the CBD	• surface (m2) × log(sizeh) • is housev(dummy) × size2h • high educi (%) × high educh • high inci (%) × mid/high incomeh • low inci (%) × high incomeh • PT accessibilityi • car accessi • industryi (jobs/m2) × high incomeh • officei (jobs/m2) × workers	• number of rooms • Age of the house • Total area of the lot • Garage • Median zonal income • Distance to downtown • Environmental attributes • Race • income
Alternatives	Aggregated households based on their age and family status	Different income groups	Different income groups	Income, presence or absence of children

Appendix D: Visualized accessibility measurements

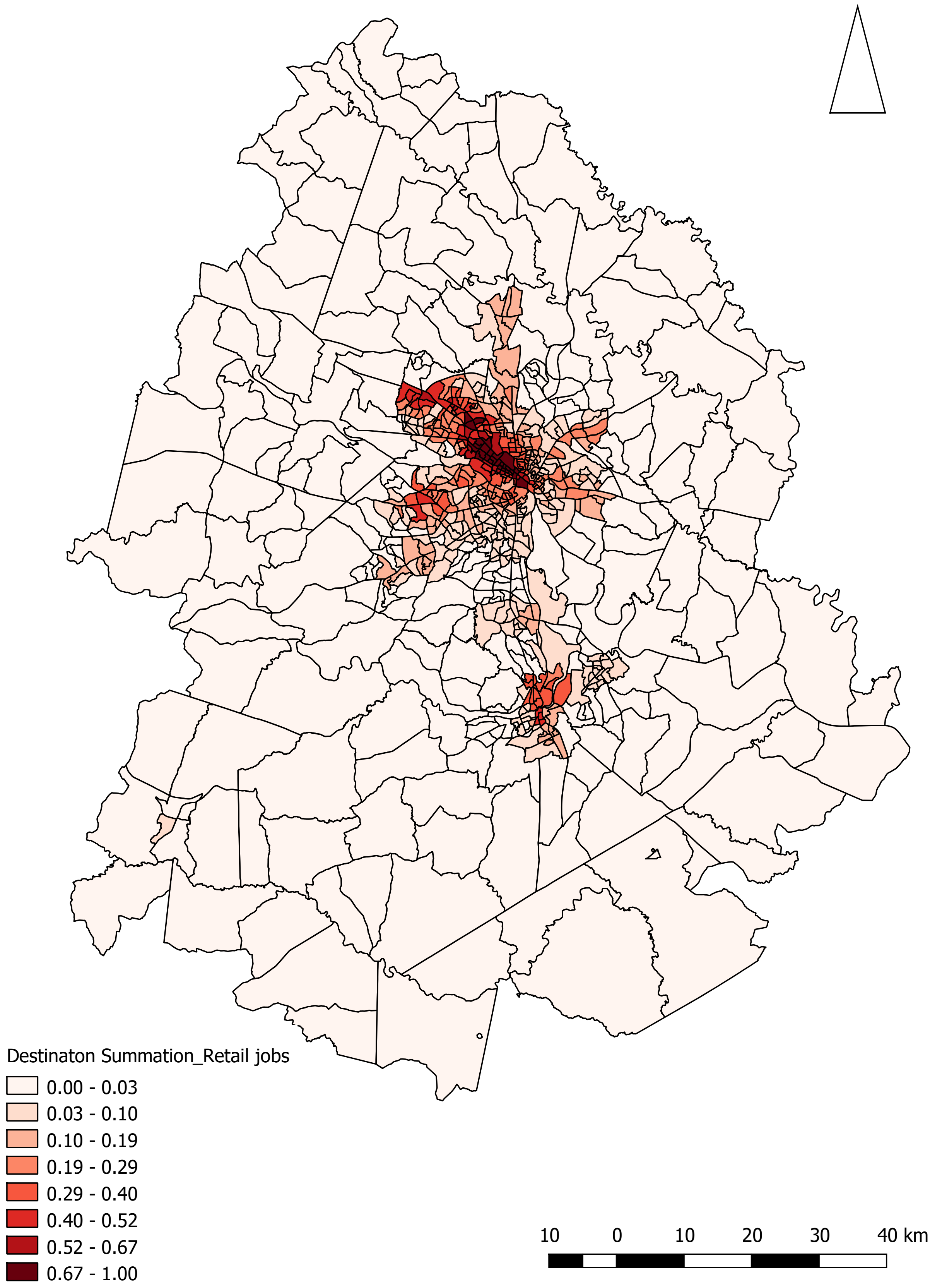

Fig. 27: Traditional cumulative accessibility measurements (Destination Summation) for retail jobs

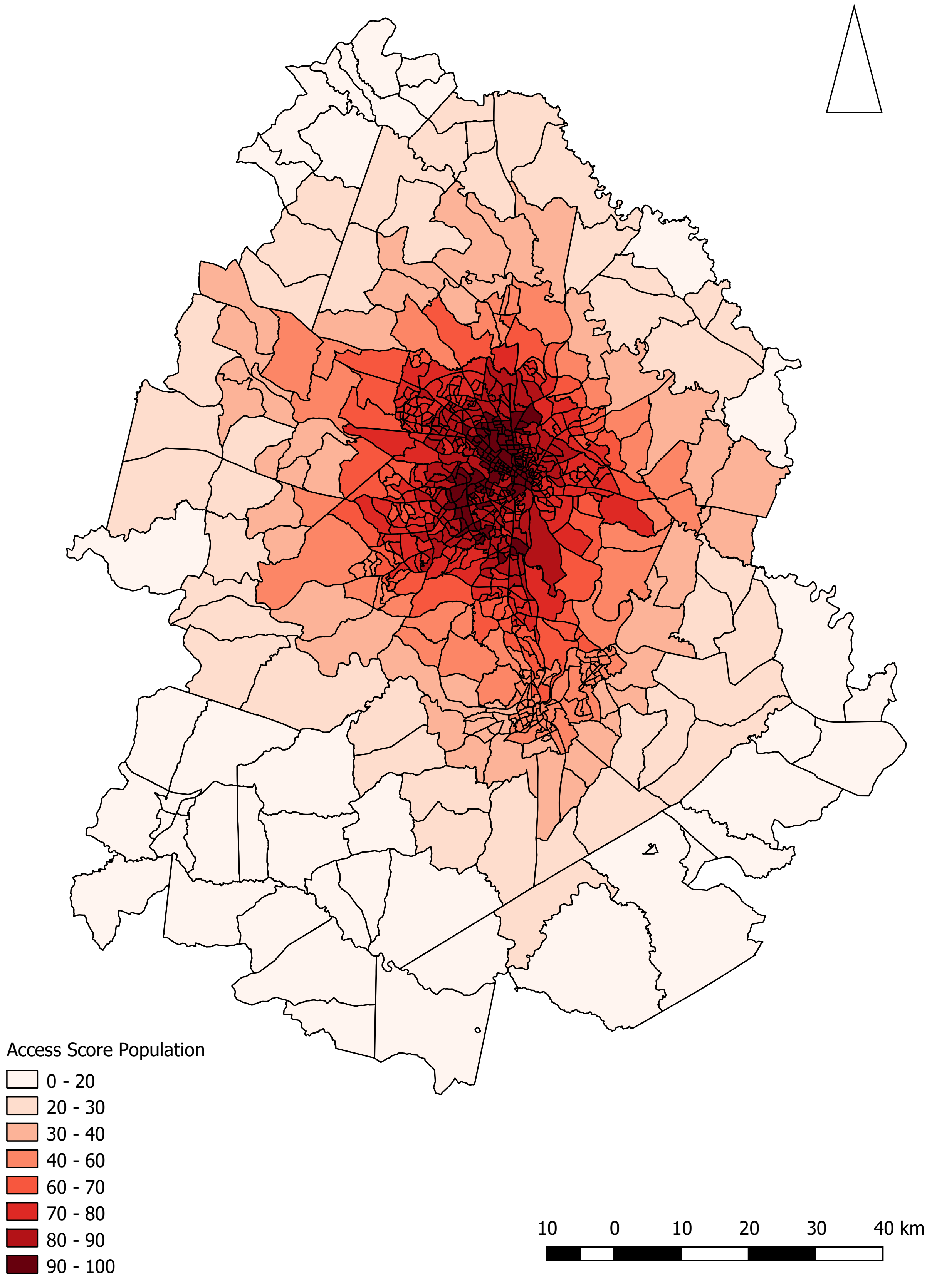

Fig. 28: Cumulative accessibility measurement (Access Score) for population

Appendix E: The Biogeme output of aggregation

Table 12: Statistical summary of different aggregation estimations

Parameter	3 income groups	6 income groups
Model	Logit	Logit
Number of observations	2010	2010
Final log likelihood	-1977.246	-3201.283
Rho-square	0.105	0.111
Accessibility measurement	Access score	Access score

Output of 3 income groups:

Model:	Logit
Number of estimated parameters:	11
Number of observations:	2010
Null log likelihood:	-2208.211
Cte log likelihood:	-1995.014
Init log likelihood:	-2136.846
Final log likelihood:	-1977.246
Likelihood ratio test:	461.930
Rho-square:	0.105
Final gradient norm:	+8.175e-003
Diagnostic:	Convergence reached...
Variance-covariance:	from analytical hessian

Name	Value	Std err	t-test	p-value
ASC_Income1	-0.0517	0.190	-0.27	0.79
ASC_Income2	0.138	0.163	0.85	0.40
ASC_Income3	0.000	fixed		
BETA_AccessScoreWorkCar1	0.00872	0.00233	3.74	0.00
BETA_AccessScoreWorkCar3	0.000	fixed		
BETA_Child1	-0.459	0.158	-2.91	0.00
BETA_Child2	-0.662	0.132	-5.01	0.00
BETA_Child3	0.000	fixed		
BETA_Gender1	-1.23	0.121	-10.18	0.00
BETA_Gender2	-0.502	0.111	-4.52	0.00
BETA_Gender3	0.000	fixed		
BETA_HomeTypeHhsize2	0.0607	0.0307	1.97	0.05
BETA_HomeTypeHhsize3	0.000	fixed		
BETA_Retired1	1.30	0.143	9.08	0.00
BETA_Retired2	0.404	0.132	3.07	0.00
BETA_Retired3	0.000	fixed		
BETA_URBRUR2	0.329	0.112	2.93	0.00
BETA_URBRUR3	0.000	fixed		

Name	Value	Std err	t-test	p-value
Lnhh1	11.6	fixed		
Lnhh2	11.8	fixed		
Lnhh3	12.3	fixed		

Output of 6 income groups:

Model: Logit
Number of estimated parameters: 20
Number of observations: 2010
Null log likelihood: -3601.437
Cte log likelihood: -3145.937
Init log likelihood: -3383.951
Final log likelihood: -3201.283
Likelihood ratio test: 800.307
Rho-square: 0.111
Final gradient norm: +1.286e-002
Diagnostic: Convergence reached...
Variance-covariance: from analytical hessian

Name	Value	Std err	t-test	p-value
ASC_Income1	-0.364	0.272	-1.33	0.18
ASC_Income2	0.808	0.140	5.75	0.00
ASC_Income3	0.544	0.141	3.85	0.00
ASC_Income4	0.653	0.165	3.97	0.00
ASC_Income5	0.546	0.140	3.89	0.00
ASC_Income6	0.000	fixed		
BETA_AccessScoreWork1	0.0151	0.00334	4.52	0.00
BETA_AccessScoreWork6	0.000	fixed		
BETA_Child2	-0.845	0.197	-4.29	0.00
BETA_Child3	-0.628	0.175	-3.59	0.00
BETA_Child4	-0.901	0.153	-5.91	0.00
BETA_Child5	-0.403	0.164	-2.46	0.01
BETA_Child6	0.000	fixed		
BETA_Gender1	-1.29	0.166	-7.76	0.00
BETA_Gender2	-1.48	0.151	-9.80	0.00
BETA_Gender3	-0.798	0.147	-5.44	0.00
BETA_Gender4	-0.484	0.144	-3.35	0.00
BETA_Gender5	-0.588	0.164	-3.60	0.00
BETA_Gender6	0.000	fixed		
BETA_Retired1	1.52	0.158	9.64	0.00
BETA_Retired2	1.11	0.155	7.16	0.00
BETA_Retired3	0.565	0.156	3.61	0.00
BETA_Retired6	0.000	fixed		

Name	Value	Std err	t-test	p-value
BETA_URBRUR1	-0.426	0.154	-2.76	0.01
BETA_URBRUR4	0.265	0.147	1.81	0.07
BETA_URBRUR6	0.000	fixed		
lnhh1	10.7	fixed		
lnhh2	11.0	fixed		
lnhh3	11.1	fixed		
lnhh4	11.0	fixed		
lnhh5	10.8	fixed		
lnhh6	12.0	fixed		

Appendix F: The Biogeme output of traditional cumulative accessibility measurement to retail jobs

Model:	Logit
Number of estimated parameters:	12
Number of observations:	2010
Number of individuals:	2010
Null log likelihood:	-2208.211
Cte log likelihood:	-1995.014
Init log likelihood:	-2136.846
Final log likelihood:	-1941.270
Likelihood ratio test:	533.882
Rho-square:	0.121
Adjusted rho-square:	0.115
Final gradient norm:	+1.314e-002
Diagnostic:	Convergence reached...
Iterations:	14
Variance-covariance:	from analytical hessian

Name	Value	Std err	t-test	p-value
ASC_Income1	-0.107	0.137	-0.78	0.43
ASC_Income2	-0.0207	0.170	-0.12	0.90
ASC_Income3	0.000	fixed		
BETA_AutoDesRetail1	3.35	0.422	7.94	0.00
BETA_AutoDesRetail2	1.04	0.427	2.44	0.01
BETA_AutoDesRetail3	0.000	fixed		
BETA_Child1	-0.329	0.164	-2.01	0.04
BETA_Child2	-0.617	0.132	-4.66	0.00
BETA_Child3	0.000	fixed		
BETA_Gender1	-1.19	0.124	-9.61	0.00
BETA_Gender2	-0.474	0.110	-4.30	0.00
BETA_Gender3	0.000	fixed		
BETA_HomeTypeHhsize2	0.0627	0.0307	2.04	0.04
BETA_HomeTypeHhsize3	0.000	fixed		
BETA_Retired1	1.43	0.148	9.64	0.00
BETA_Retired2	0.438	0.131	3.34	0.00
BETA_Retired3	0.000	fixed		
BETA_URBRUR2	0.317	0.112	2.83	0.00
BETA_URBRUR3	0.000	fixed		
Lnhh1	11.6	fixed		
Lnhh2	11.8	fixed		
Lnhh3	12.3	fixed		

Appendix G: The Biogeme output of gravity Eq. 2-4 accessibility measurements

$\alpha = 1.0, \beta = 0.5$:

Model:	Logit
Number of estimated parameters:	12
Number of observations:	2010
Null log likelihood:	-2208.211
Cte log likelihood:	-1995.014
Init log likelihood:	-2136.846
Final log likelihood:	-1963.762
Likelihood ratio test:	488.897
Rho-square:	0.111
Diagnostic:	Convergence reached...
Variance-covariance:	from analytical hessian

Name	Value	Std err	t-test	p-value
ASC_Income1	-0.0191	0.141	-0.14	0.89
ASC_Income2	-0.0758	0.175	-0.43	0.66
ASC_Income3	0.000	fixed		
BETA_AutoHansenPop1	3.21	0.513	6.26	0.00
BETA_AutoHansenPop2	1.40	0.481	2.91	0.00
BETA_AutoHansenPop3	0.000	fixed		
BETA_Child1	-0.442	0.160	-2.77	0.01
BETA_Child2	-0.630	0.132	-4.76	0.00
BETA_Child3	0.000	fixed		
BETA_Gender1	-1.23	0.122	-10.04	0.00
BETA_Gender2	-0.478	0.111	-4.30	0.00
BETA_Gender3	0.000	fixed		
BETA_HomeTypeHhsize2	0.0588	0.0308	1.91	0.06
BETA_HomeTypeHhsize3	0.000	fixed		
BETA_Retired1	1.35	0.145	9.34	0.00
BETA_Retired2	0.438	0.132	3.32	0.00
BETA_Retired3	0.000	fixed		
BETA_URBRUR2	0.321	0.112	2.85	0.00
BETA_URBRUR3	0.000	fixed		
Lnhh1	11.6	fixed		
Lnhh2	11.8	fixed		
Lnhh3	12.3	fixed		

$\boldsymbol{\alpha = 1.0, \beta = 0.3}$:

Model:	Logit
Number of estimated parameters:	12
Number of observations:	2010
Null log likelihood:	-2208.211
Cte log likelihood:	-1995.014
Init log likelihood:	-2136.846
Final log likelihood:	-1971.342
Likelihood ratio test:	473.737
Rho-square:	0.107
Final gradient norm:	+1.297e-002
Diagnostic:	Convergence reached...
Variance-covariance:	from analytical hessian

Name	**Value**	**Std err**	**t-test**	**p-value**
ASC_Income1	0.0799	0.142	0.56	0.57
ASC_Income2	-0.00337	0.176	-0.02	0.98
ASC_Income3	0.000	fixed		
BETA_Child1	-0.468	0.158	-2.95	0.00
BETA_Child2	-0.645	0.132	-4.87	0.00
BETA_Child3	0.000	fixed		
BETA_Gender1	-1.24	0.122	-10.21	0.00
BETA_Gender2	-0.488	0.111	-4.39	0.00
BETA_Gender3	0.000	fixed		
BETA_HomeTypeHhsize2	0.0594	0.0308	1.93	0.05
BETA_HomeTypeHhsize3	0.000	fixed		
BETA_Htwo1	2.66	0.526	5.05	0.00
BETA_Htwo2	0.917	0.497	1.85	0.06
BETA_Htwo3	0.000	fixed		
BETA_Retired1	1.32	0.144	9.19	0.00
BETA_Retired2	0.422	0.132	3.20	0.00
BETA_Retired3	0.000	fixed		
BETA_URBRUR2	0.323	0.112	2.88	0.00
BETA_URBRUR3	0.000	fixed		
Lnhh1	11.6	fixed		
Lnhh2	11.8	fixed		
Lnhh3	12.3	fixed		

$\boldsymbol{\alpha = 1.5, \beta = 0.5}$:

Model:	Logit
Number of estimated parameters:	11
Number of observations:	2010
Null log likelihood:	-2208.211
Cte log likelihood:	-1995.014
Init log likelihood:	-2136.846
Final log likelihood:	-1977.364
Likelihood ratio test:	461.694
Rho-square:	0.105
Final gradient norm:	+9.682e-003
Diagnostic:	Convergence reached...
Variance-covariance:	from analytical hessian

Name	Value	Std err	t-test	p-value
ASC_Income1	0.292	0.131	2.23	0.03
ASC_Income2	0.138	0.163	0.85	0.40
ASC_Income3	0.000	fixed		
BETA_Child1	-0.496	0.158	-3.14	0.00
BETA_Child2	-0.663	0.132	-5.02	0.00
BETA_Child3	0.000	fixed		
BETA_Gender1	-1.26	0.121	-10.39	0.00
BETA_Gender2	-0.501	0.111	-4.50	0.00
BETA_Gender3	0.000	fixed		
BETA_HomeTypeHhsize2	0.0603	0.0307	1.96	0.05
BETA_HomeTypeHhsize3	0.000	fixed		
BETA_Hthree1	1.36	0.353	3.86	0.00
BETA_Hthree3	0.000	fixed		
BETA_Retired1	1.29	0.143	9.05	0.00
BETA_Retired2	0.404	0.132	3.06	0.00
BETA_Retired3	0.000	fixed		
BETA_URBRUR2	0.329	0.112	2.93	0.00
BETA_URBRUR3	0.000	fixed		
Lnhh1	11.6	fixed		
Lnhh2	11.8	fixed		
Lnhh3	12.3	fixed		

$\alpha = 1.5, \beta = 0.3$:

Model:	Logit
Number of estimated parameters:	11
Number of observations:	2010
Null log likelihood:	-2208.211
Cte log likelihood:	-1995.014
Init log likelihood:	-2136.846
Final log likelihood:	-1981.204
Likelihood ratio test:	454.013
Rho-square:	0.103
Final gradient norm:	+8.178e-003
Diagnostic:	Convergence reached...
Variance-covariance:	from analytical hessian

Name	Value	Std err	t-test	p-value
ASC_Income1	0.384	0.129	2.99	0.00
ASC_Income2	0.140	0.163	0.86	0.39
ASC_Income3	0.000	fixed		
BETA_Child1	-0.514	0.157	-3.27	0.00
BETA_Child2	-0.663	0.132	-5.01	0.00
BETA_Child3	0.000	fixed		
BETA_Gender1	-1.27	0.121	-10.49	0.00
BETA_Gender2	-0.501	0.111	-4.50	0.00
BETA_Gender3	0.000	fixed		
BETA_Hfour1	0.881	0.328	2.68	0.01
BETA_Hfour3	0.000	fixed		
BETA_HomeTypeHhsize2	0.0600	0.0308	1.95	0.05
BETA_HomeTypeHhsize3	0.000	fixed		
BETA_Retired1	1.28	0.143	8.96	0.00
BETA_Retired2	0.403	0.132	3.06	0.00
BETA_Retired3	0.000	fixed		
BETA_URBRUR2	0.328	0.112	2.92	0.00
BETA_URBRUR3	0.000	fixed		
Lnhh1	11.6	fixed		
Lnhh2	11.8	fixed		
Lnhh3	12.3	fixed		