Baraklianos, Ioannis

Thesis

The accessibility in Land-Use Transport Interaction models : four essays on location choice models

The accessibility in Land-Use Transport Interaction models: four essays on location choice models

Ioannis Baraklianos

N° d'ordre NNT : 2019LYSE2037

THESE de DOCTORAT DE L'UNIVERSITÉ DE LYON

Opérée au sein de

L'UNIVERSITÉ LUMIÈRE LYON 2

École Doctorale : ED 486 Sciences Économique et de Gestion

Discipline : Sciences économique

Soutenue publiquement le 26 juin 2019, par :

Ioannis BARAKLIANOS

The accessibility in Land-Use Transport Interaction Models.

Four essays on location choice models.

Devant le jury composé de :

Michel DIMOU, Professeur des universités, Université de Toulon, Président

Catherine BAUMONT, Professeure des universités, Université de Bourgogne, Rapporteure

Faridah DJELLAL, Professeure des universités, Université de Lille, Rapporteure

Yves CROISSANT, Professeur des universités, Université de La Réunion, Examinateur

Patrick BONNEL, Ingénieur de Recherche, École Nationale des Travaux Publics de l'État, Co-Directeur de thèse

Louafi BOUZOUINA, Chercheur, École Nationale des Travaux Publics de l'État, Co-Directeur de thèse

Contrat de diffusion

Thèse de Doctorat de l'Université Lumière Lyon 2

Préparée à l'Ecole Nationale des Travaux Publics de l'Etat

Faculté de Sciences Économiques et de Gestion
Ecole Doctoral 486

The accessibility in Land-Use Transport Interaction models.
Four essays on location choice models.

Soutenue publiquement le 24 juin 2019 par

Ioannis BARAKLIANOS

Devant le jury compose de :

Michel DIMOU	Université de Toulon	Professeur des universités	Président
Catherine BAUMONT	Université de Bourgogne	Professeur des universités	Rapportrice
Faridah DJELLAL	Université de Lille	Professeur des universités	Rapportrice
Yves CROISSANT	Université de la Réunion	Professeur des universités	Examinateur
Louafi BOUZOUINA	Ecole Nationale des Travaux Publics de l'Etat	Chercheur	Co-directeur de thèse
Patrick BONNEL	Ecole Nationale des Travaux Publics de l'Etat	Chercheur, ingénieur en chef	Co-directeur de thèse

Thèse de Doctorat de l'Université Lumière Lyon 2

Préparée à l'Ecole Nationale des Travaux Publics de l'Etat

Faculté de Sciences Économiques et de Gestion
Ecole Doctoral 486

The accessibility in Land-Use Transport Interaction models.

Four essays on location choice models.

Soutenue publiquement le 24 juin 2019 par

Ioannis BARAKLIANOS

Devant le jury compose de :

Michel DIMOU	Université de Toulon	Professeur des universités	Président
Catherine BAUMONT	Université de Bourgogne	Professeur des universités	Rapportrice
Faridah DJELLAL	Université de Lille	Professeur des universités	Rapportrice
Yves CROISSANT	Université de la Réunion	Professeur des universités	Examinateur
Louafi BOUZOUINA	Ecole Nationale des Travaux Publics de l'Etat	Chercheur	Co-directeur de thèse
Patrick BONNEL	Ecole Nationale des Travaux Publics de l'Etat	Chercheur, ingénieur en chef	Co-directeur de thèse

« Της παιδείας οι μεν ρίζες πικρές, οι δε καρποί γλυκείς »
Αριστοτέλης (384 – 322 π.Χ.)

« The roots of the education are bitter, but the fruits are sweet »
Aristotle (384 – 322 B.C.)

« Les racines de l'éducation sont amères, mais les fruits en sont doux »
Aristote (384 – 322 av. J.-C.)

Acknowledgements

While a PhD thesis is a personal work, without the valuable help of the following I would not be able to arrive at the end.

First of all, I would like to express my gratitude to my thesis advisors Louafi Bouzouina and Patrick Bonnel. Throughout these almost four years, they have always been supportive in all levels with their experience, recommendations, ideas, understanding and encouragement. If it weren't for them, I wouldn't be able to finish my thesis in such good terms.

I also thank Prof. Catherine Baumont and Prof. Faridah Djellal for accepting to be the rapporteurs of this thesis and Prof. Michel Dimou and Prof. Yves Croissant for agreeing to participate to the jury.

I own a special thanks to my co-authors of some articles, Hind Aissaoui for helping me at the beginning of the thesis and for providing me valuable advices and data that saved me time and Ouassim Manout with which I had the opportunity to collaborate in various instances.

A big thank you to all of my colleagues of the LAET with which I shared special moments in and out of the laboratory.

I thank the Region of Auvergne-Rhône-Alpes for the financial support, which permitted me to fully concentrate my time to this thesis.

I thank my friend Tasos, who passed many hours reading and correcting parts of the manuscript.

A special thanks to my parents and my sister who helped me realise my project despite the distance and difficulties. I thank also all my friends in Greece and elsewhere for their presence and support.

I reserve my biggest thank you to my "*co-équipier*" Ricci and my boys for believing in me and for their constant presence, support and understanding during this challenging period.

Notice

The four chapters of this dissertation are independent research articles (see list below). This explains why some information can be redundant, and why the terms paper or article are used. The main responsible of these articles is Ioannis Baraklianos. Louafi Bouzouina and Patrick Bonnel accompanied the conception and the redaction of each article with their expertise and experience. The first paper/chapter of this dissertation is co-authored also with Hind Aissaoui and is published in “Transportation” journal (Baraklianos et al., 2018b). Hind Aissaoui provided data and the R script for the estimation of the models. The second paper/chapter is co-authored also with Ouassim Manout and is accepted for publication by the journal “Papers in Regional Science”. Ouassim Manout provided generalised times for the estimation of the accessibility indicators. The third paper/chapter is published in “Region and Development” journal (Baraklianos et al., 2018a). The fourth paper/chapter is co-authored also with Hind Aissaoui and is a work in progress. Hind Aissaoui provided data for the estimation of the models.

1. Baraklianos, I., Bouzouina, L., Bonnel, P., & Aissaoui, H. (2018). Does the accessibility measure influence the results of residential location choice modelling? *Transportation*. https://doi.org/s11116-018-9964-6
2. Baraklianos, I., Bouzouina, L., Bonnel, P., & Manout, O. (2019). Do new and relocating firms have different preferences for accessibility? *Papers in Regional Science*, *accepted*.
3. Baraklianos, I., Bouzouina, L., & Bonnel, P. (2018). The impact of accessibility on the location choices of the business services. Evidence from Lyon urban area. *Région et Développement*, *48*, 85–104.
4. Baraklianos, I., Bouzouina, L., Bonnel, P., & Aissaoui, H. Renters vs Owners. How has the preference for accessibility evolved for a residential location choice? The case of the Lyon urban area in France, 1999-2013. *Working paper*

Table of Contents

Introduction 11
The concept of accessibility 11
Historical evolution 12
Components and dimensions 14
LUTI models 15
Definition and objective 15
Structure 16
The SIMBAD model 18
The key role of accessibility in the modelling chain 19
Integrating accessibility in location choice models 21
Accessibility measurements 21
Identified challenges 23
Four essays on accessibility in location choice models 25
Does the accessibility measure influence the results of residential location choice modelling? 26
Do new and relocating firms have different preferences for accessibility? 27
What is the impact of accessibility on the location choices of business services? 27
Renters vs owners. How has the preference for accessibility evolved in residential location choices 27
Transversal choices 28
Chapter I Does the accessibility measure influence the results of residential location choice modelling? 32

1. Introduction 32

2. Accessibility and residential location choice: one concept, different approaches, different results? 34

3. Methodology and application 40

3.1. Residential location choice model 40

3.2. Accessibility measures and implementation 41

3.3. Study area and data 44

3.4. Model variables and measures 46

3.5. Evaluation method 47

4. Results 49

4.1. Analysis of the model parameters 49

4.2. Model quality and statistical contribution of accessibility 50

4.3. Ability to replicate market shares 52

5. Summary and discussion 54

Chapter II Do new and relocating firms have different preferences for accessibility? 58

1. Introduction 58

2. Differences between new and relocating establishments: do they appreciate accessibility differently? 59

3. Study area, data and model specification 61

3.1. The Urban Area of Lyon 61

3.2. Creations and migrations of economic establishments 63

3.3. Modelling the location choices of economic establishments 66

3.4 The independent variables 68

4. Modelling results 73

4.1. New and relocating economic establishments choose their locations differently 74

4.2. Different events, different preferences 75

5. Conclusions 79

6. Appendix 83

Chapter III The impact of accessibility on the location choices of business services. 87

1. Introduction 87

2. Study area 90

3. Data sources 92

4. Modelling the location choices 93

   4.1. Model specification 93

   4.2. Variables and measures 94

5. Location choices of Front Office and Back Office business services – descriptive statistics 98

6. Modelling results - Front Office, Back Office services: Diverging choices 101

7. Conclusions and perspectives 107

Chapter IV Renters vs owners. How has the preference for accessibility evolved in residential location choices? 110

1. Introduction 110

2. Related empirical studies 112

3. Study area 114

   3.1. Evolution of the population 114

   3.2 Evolution of public transport supply and accessibility 116

4. Method and data 117

   4.1. Residential location choice model 117

   4.2 Determinants of location choices 120

5. Results and analysis 122

6. Conclusions 126

7. Appendix 129

Conclusion 132

PhD thesis contributions 135

Does the accessibility measure influence the results of residential location choice models? 135

Do new and relocating firms have different preferences for accessibility? 137

What is the impact of accessibility on the location choices of the business services? ... 139

Renters vs owners. How has the preference for accessibility evolved in residential location choices? 141

General contribution 143

Limits and perspectives 144

Bibliography 146

Abstract 160

Résumé 162

List of figures

Figure 1 Components and dimensions of accessibility, adapted from Geurs and Wee, 2004 14

Figure 2 The Land-use – Transport feed-back loop, adapted from Wegener and Fürst, 1999.............17

Figure 3 General structure of LUTI model and PhD thesis focus 17

Figure 4 The structure of the SIMBAD model of the land-use and transport sub-models 18

Figure 5 The key role of the accessibility in LUTI models 20

Figure 6 The balance of an accessibility measure choice 24

Figure 7 Relation of the PhD thesis papers to the research axes 26

Figure I.1 Employment density of the study area by zone (classification in quantiles) 45

Figure I.2 The zones of the urban area retained for the RMSE analysis 49

Figure II.1 The zones and the transport infrastructure of the urban area of Lyon. 62

Figure III.1 Study area and transport infrastructure 91

Figure III.2 Preference for accessibility to population of new and migrating firms 99

Figure III.3 Location areas of new and migrating firms 99

Figure III.4 Location choices of new Front Office establishments 100

Figure III.5 Location choices of new Back Office establishments 100

Figure III.6 Location choices of relocating Back Office establishments 100

Figure III.7 Location choices of relocation Front Office establishments 100

Figure IV.1 Evolution of population from 1968 to 2013 by analysis zone 115

Figure IV.2 Evolution of the number of households from 1999 to 2013 by aggregated analysis zone (figure IV.3) and housing ownership status 115

Figure. IV.3 Aggregate urban area zoning and the metro/tramway network evolution 117

Figure IV.4 Evolution of elasticities by variable group and by renters/owners Source: Authors' estimations 125

Figure IV.A1. Evolution of elasticities by variable group and by renters/owners Source: Authors' estimations 129

List of tables

Table 1 Evolution of accessibility over time 12

Table I.1 Literature review summary 39

Table I.2 Correlation (spatial correlation - global Moran's I) of the number of jobs per zone 44

Table I.3 Distribution of recently moved households' characteristics 46

Table I.4 Descriptive statistics of the zonal attribitues of the variables model 47

Table I.5 Parameters and statistical tests of the models using all accessibility measures 51

Table I.6 Parameters and statistical tests of the models including interaction terms 52

Table I.7 Descriptive statistics of the accessibility variables for the aggregated zones 53

Table I.8 Capacity to replicate the observed market share of the models using all accessibility measures 54

Table I.9 Capacity to replicate the observed market share of the models including systematic taste variation and potential accessibility measures 54

Table II.1 Classification of economic establishments by INSEE and modifications 64

Table II.2 Creation and relocation of establishments by sector 65

Table II.3 Distribution of the establishments' size included into the analysis by sector and creation/relocation 65

Table II.4 Summary statistics of the variables used in the model 72

Table II.5 Likelihood ratio test for segmentation between created and relocated establishments 74

Table II.6 Elasticities for creations and migrations in Manufacturing and Construction 78

Table II.7 Elasticities for creations and migrations in Wholesale and Retail 78

Table II.8 Elasticities for creations and migrations in R.E Finance & Insurance and Front Office 79

Table II.9 Elasticities for creations and migrations in Back Office and Health 79

Table II.10 Synthesis of the observed differences in preferences for accessibility for all types of accessibility and all sectors 81

Table II.A1 Model results of manufacturing, construction and wholesale sectors 83

Table II.A2 Model results of retail, R.E Finance & Insurance and Front Office sectors 84

Table II.A3 Model results of back office and health sectors 85

Table III.1 Key statistics of creations and relocations in Front Office and Back Office services 99

Table III.2 Estimated parameters of the location choices of Front Office and Back Office services ..103

Table III.3 Estimated mean elasticities for Front Office and Back Office services...........................104

Table III.4 Results of the t-test for the variables of the Front Office versus Back Office services 106

Table IV.1 Accessibility evolution by aggregate zoning zones 116

Table IV.2 Distribution of recently moved households by aggregate zoning and housing occupancy status...120

Table IV.3 Descriptive statistics and spatial autocorrelation of the zonal attributes 122

Table IV.4. Estimated parameters and significance of the models Source: Authors' calculations.....123

Table IV.5. Car ownership based on the housing occupancy status of the recently moved households 126

Table I.VA1. Distribution of authorised surface for housing by aggregated zone 130

Introduction

Accessibility is by definition the interaction between transport and land use. Hansen (1959 p.73) was the first to give a formal definition. He defined accessibility as the "*potential of opportunities for interaction*". Since then, many have extended what accessibility means and how it is measured. Today, accessibility is considered to be the main service offered by a transport system to its users (Geurs and Ritsema van Eck, 2001). It is not only one of the most extensively used indicators to evaluate urban planning policy efficiency (Handy and Niemeier, 1997) but it is also the key construction element of most Land-Use Transport Interaction (LUTI) models (Acheampong and Silva, 2015; Baraklianos et al., 2018b; Wegener and Fürst, 1999). Accessibility is in the heart of LUTI models. It translates any transport or land-use change into a quantifiable amount, which is integrated into the LUTI modelling chain and allows the continuous interaction between the different components of a LUTI model.

While accessibility is a very useful and powerful tool, it still remains an inherently vague concept. It is something that we cannot measure directly, like the number of cars passing through an intersection or the population living in a specific area. Accessibility is a concept that we need to model. Consequently, the modelling procedure implicates simplifications of the reality and methodological choices (Bonnel, 2004). In the beginning, this modelling procedure was empirical and with no theoretical framework (Geurs and van Wee, 2004). The theoretical bases came later, after proving its value as an analysis tool. Each discipline (economy, geography, phycology, etc.) embraced the generality and adaptability of the accessibility concept and gave its own theoretical bases and specific definition of what accessibility is, having a different focus and using different assumptions (Baraklianos et al., 2018b). This "flexibility" of the accessibility concept is its strength and its weakness at the same time.

LUTI models are no exception to the fundamental challenges of accessibility. Despite the rich literature on the subject, there are still a lot of LUTI–accessibility unanswered questions. Accessibility has a key role in LUTI models. It assures the interaction between the two main components of the LUTI models, the transport and land-use sub-models and is present in different modules (location choice models, relocation models, hedonic price model, etc.). Between the different modules of LUTI models, in this thesis we focus on the location choice models. Their interest does not only concern model scientists, but also policy design. The location choices of households and firms are the two reference points explaining everyday mobility (Acheampong and Silva, 2015). Issues related to the location choices of households and firms are at the centre of any integrated transport–land-use policy (Homocianu, 2009). This thesis examines the place and the importance of accessibility in the location choice models of households and firms. The objective is to analyse the effect of various methodological choices from a theoretical and empirical point of view and to give some answers to methodological and policy issues. This broad problematic is specified in four research papers, which constitute the main part of this PhD thesis.

The rest of this introductory section is structured as follows. Section 1.1 presents the historical evolution of the accessibility concept and its different components and dimensions. Section 1.2 briefly presents the logic of LUTI models and explains why accessibility is a key element of LUTI models and its role in the location choice models. Section 1.3 presents how accessibility is operationalised and the identified challenges of the application of the accessibility in the context of location choice models. Last, section 1.4 presents the four research papers, which constitute the main work of this thesis, and the shared methodological choices.

The concept of accessibility

Historical evolution

Accessibility is an inherently vague concept for academia because it is first an empirical tool. The theoretical bases came after its application in various empirical studies. As it is a concept that we model, it is difficult to give a clear, theoretically sound and generally accepted definition. This is why Geurs & van Wee (2004, p. 127) point out that "*accessibility is often a misunderstood, poorly defined and poorly measured construct*". If we look at some selected definitions of accessibility given in the literature throughout time (table 1), we can understand that the accessibility concept is something that evolves. Researchers integrate more and more elements into its definition, thanks to advances in different fields and technological innovations, in order to be as complete as possible. Yet, because accessibility is first an empirical tool, researchers need to define how they model, integrate and interpret accessibility into their works.

***Table 1** Evolution of accessibility definition over time*

Author	**Definition**
Hansen, 1959	"…is the potential of opportunities for interaction"
Dalvi & Martin, 1976	"…indicates the inherent characteristic (or advantage) of a place with respect to overcoming some form of spatially operating source of friction"
Ben-Akiva & Lerman, 1979	"…represents the benefits provided by a transportation/land-use system"
Handy & Niemeier, 1997	"…is determined by the spatial distribution of potential destinations, the ease of reaching each destination, and the magnitude, quality, and character of the activities found there"
Geurs & van Wee, 2004	"…is the extent to which land-use and transport systems enable (groups of) individuals to reach activities or destinations by means of a (combination of) transport mode(s)"
Lättman et al., 2016	"…how easy it is to live a satisfactory life using the transport system"

Today, it is generally accepted that accessibility is a construct of different components and dimensions (figure 1). However, accessibility was not conceived as a multicomponent and multidimensional concept at its birth. The concept evolved together with the understanding of the interactions between land use and transport. The advances in various fields, notably in economics and geography, influenced the concept of accessibility. Initially, it was merely a two-component concept (Hansen, 1959), defined as the interaction between the component of transport and the component of land use (Baraklianos et al., 2018b). Alonso (1964) integrated accessibility in his works as a simple linear distance between a location and the city centre. Wilson (1971) gave the theoretical basis of gravity measures. The theories of time-travel budgets (Zahavi, 1974) and space-time prism (Hägerstrand, 1970) put the bases for the integration of the temporal dimension of accessibility. Further research and the development of the Random Utility Maximisation (RUM) theory highlighted the importance of individual components and gave the operational tool (Discrete Choice Model – DCM) for the calculation of accessibility (Ben-Akiva and Lerman, 1985). This evolution led to a fragmentation of the applied methods, which today are based on different methods with different theoretical backgrounds, methodological processes and assumptions, and different data needs.

Components and dimensions

Accessibility is a powerful tool for land-use–transport policy evaluation. Its wide application is based on some reasonable assumptions regarding individual mobility and how we value the supply of various activities. The economic interpretation of accessibility is based on three main assumptions (Halden, 2002; Koenig, 1974):

- Travel is an induced activity;
- Utility increases with supply of activities;
- Utility decreases the cost.

These assumptions provide the framework of causal relations. As in any modelling procedure, we need to make some simplifications and highlight certain components of reality. Based on these assumptions, the three components highlighted in the concept of accessibility are (figure 1): (i) land use, (ii) transport and (iii) the individual. These 3 components evolve in two dimensions, which are transversal for all the components, (i) the spatial and (ii) the temporal dimensions (Baraklianos et al., 2018b; Geurs and van Wee, 2004).

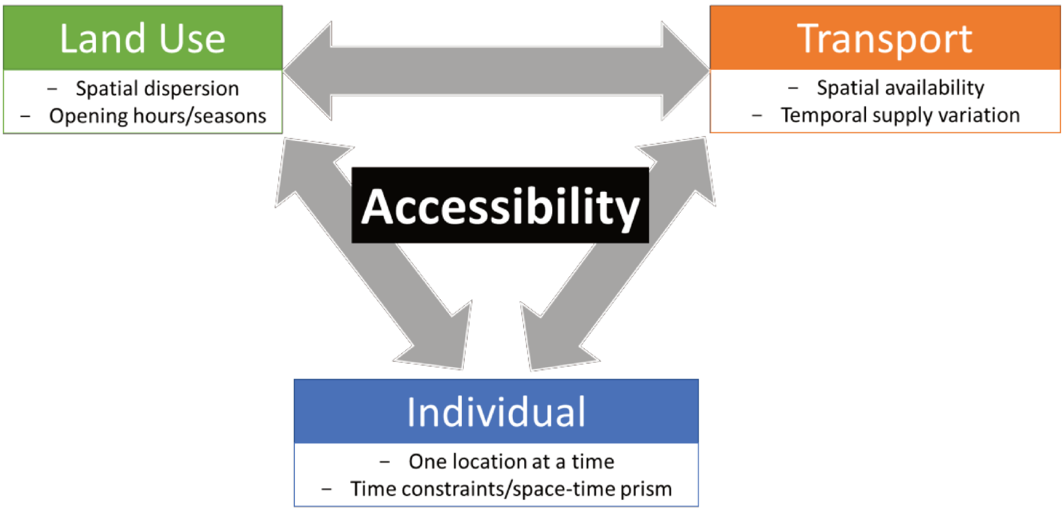


Figure 1 *Components and dimensions of accessibility, adapted from Geurs and Wee, 2004*

The **land-use** component represents the activities in which one can participate. It expresses the utilitarian part of accessibility as individuals receive utility by participating in activities (Geurs and Ritsema van Eck, 2001). The derived utility depends on the supply characteristics (amount, quality) and the balance between supply and demand, if capacity constraints exist (Geurs and van Wee, 2004). However, the activities are not equally distributed over time and space. They are spatially distributed and have specific opening hours.

The **transport component** represents transport infrastructure supply in a specific area (public transport, roads, etc.) with its characteristics (capacity, speed, etc.). Supply may differ between different times of the day or between seasons (public transport timetables, peak/off-peak, summer/winter, etc.). Under the hypothesis that travel is a derived activity (Priemus et al., 2001), the transport component represents the disutility for household members (time, cost, effort) and decreases the expected utility of an opportunity. In the accessibility framework, it is introduced as a cost that influences accessibility negatively. It decreases the utility of activities that are more difficult to be reached.

The **individual component** reflects the different characteristics (age, revenue, education level, etc.), preferences (in terms of activities, transport modes) and abilities (for example having a driving licence and owning a car) of individuals. This component can also vary in time and space. Individuals possess limited time for their needs (e.g. sleep), their obligations (e.g. work) and their other activities (e.g. leisure). Based on these time limitations, an individual can move only within a specific area, which Hägerstrand (1970) has defined as the space-time prism. Those elements affect strongly the perceived accessibility in a specific location. A location may offer great accessibility by car, but if an individual does not hold a driving licence, they cannot take advantage of this location attribute.

LUTI models

Definition and objective

LUTI models aim to represent interactions that arise in the urban environment, like mobility or the evolution of land use and urban form. Their objective is to model and forecast changes in land use and transport systems to help policy-makers take informed decisions for future policies or projects (Acheampong and Silva, 2015; Wee, 2015).

The interest for modelling the interactions between transport and land use is not new. During the last century, there have been important theoretical advances like the bid-rent theory (Alonso, 1964; Von Thünen, 1842), the entropy maximisation theory (Wilson, 1967), the RUM theory (McFadden, 1977), the time-budget theory (Zahavi, 1974), the space-time theory (Hägerstrand, 1970) or the agglomeration theory (Jacobs, 1969; Marshall, 1890). These theories were accompanied by some fundamental empirical contributions (equilibrium models, discrete choice models, spatial interaction models, gravity models) that paved the way to today's operational models. Great amount of knowledge has been cumulated in order to understand and model the complex interactions between transport and land use (Acheampong and Silva, 2015).

Even though the interest in the interactions between land use and transport dates back to the early 19th century, the first generally accepted solid works arrive during the second half of the 20th century. The arising issues of that period, like urban sprawl, congestion, pollution and pressures for land development (Weisbrod et al., 1980; Wilson, 1998), along with the technical advances, provided the appropriate circumstances to promote research around LUTI modelling.

Structure

Most of the LUTI models implement a combination of different theoretical frameworks. Therefore, each implementation can differ importantly from others (equilibrium, aggregate, utility-based, microsimulation, quasi-dynamic, dynamic). In the beginning, most of the approaches (Alonso, 1964; Lowry, 1964) were based on the principle of the equilibrium. However, an urban environment is never in equilibrium, because there are many different processes which have different reaction times (Acheampong and Silva, 2015; Wegener and Fürst, 1999). As LUTIs are models, they are based on simplifications of the urban environment and represent only certain important processes. A large number of operational LUTI models incorporate a systemic approach of the urban environment where each distinct urban process constitutes a different sub-model. The sub-models are interconnected in a predefined order and with predefined relations.

The relations in LUTI models are theoretically based on the so-called "feedback loop" (Wegener and Fürst, 1999) (figure 2). Starting from accessibility, any changes taking place in the transport system can alter the perceived accessibility. Such changes modify the attractiveness of the locations, which in turn affects the decisions of urban development projects and the location choices of households and firms. Those dynamics affect the locations of the activities, which then influence the households' mobility behaviour. These new mobility habits affect the generalised transport costs, which can provoke the need for new transport projects and therefore accessibility improvements. This dynamic is implemented into the operational LUTI models by running the modelling chain in time steps. This means that the state of the period t constitutes the departing point of the period $t+1$. The models that follow this stepwise logic are also known as quasi-dynamic models, because they are composed by many cross-sectional models for each period (Wegener and Fürst, 1999). The same relational structure is found in equilibrium models to achieve the equilibrium state.

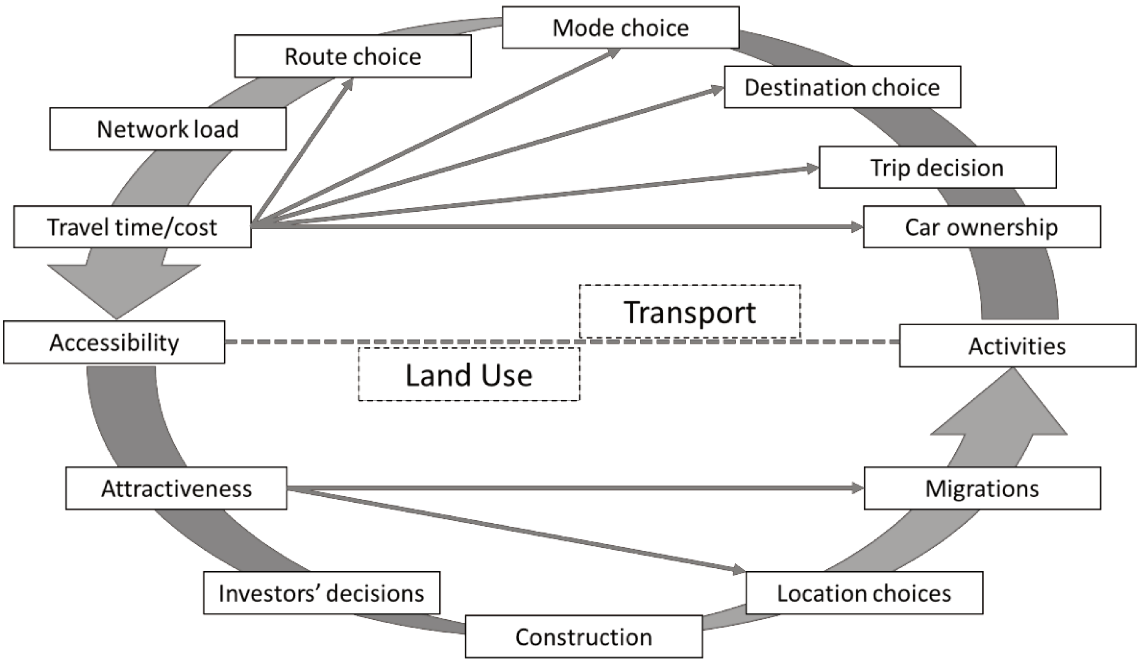


***Figure 2** The Land-use-Transport feed-back loop, adapted from Wegener and Fürst, 1999*

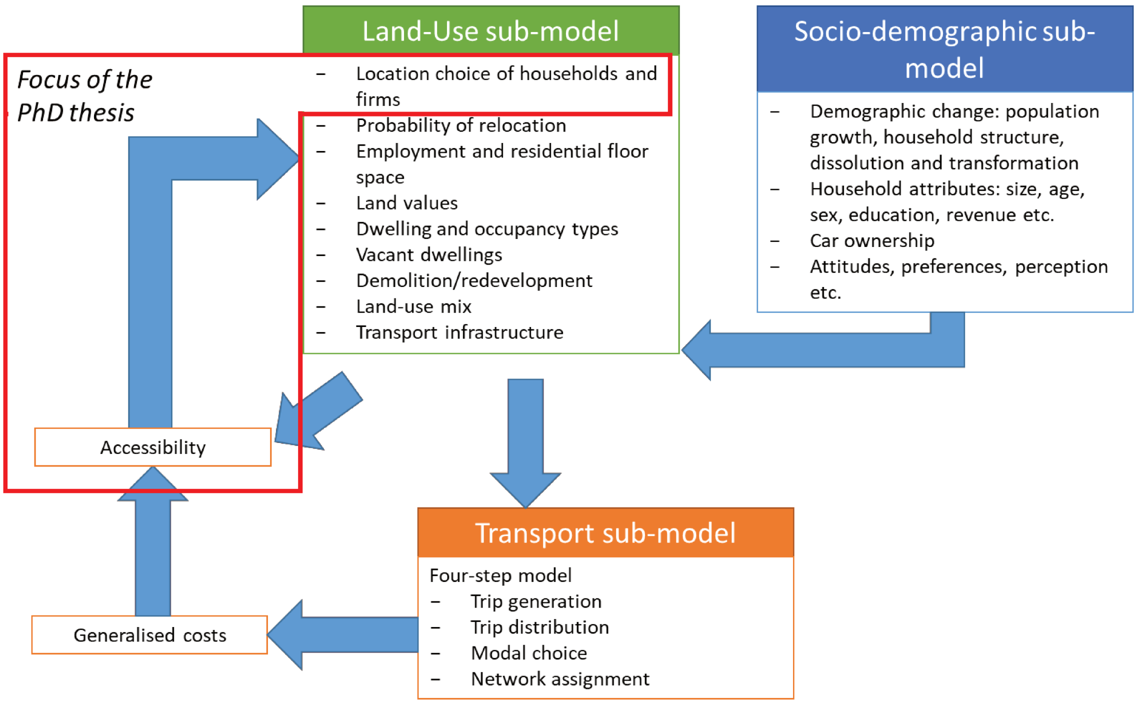


***Figure 3** General structure of LUTI model and PhD thesis focus, adapted from Acheampong and Silva, 2015*

In practice, the relational structure of the sub-models of LUTI models varies significantly between different conceptions. A general structure of operational LUTI models is presented in figure 3. The model is composed by three sub-models, each aiming to simulate a different urban process. The socio-demographic sub-model simulates the different socio-demographic characteristics of the population. The land-use sub-model simulates the processes and the interrelations of the locations. Lastly, the transport model simulates the mobility behaviour. In this PhD thesis, the focus is placed on accessibility measures and their implementation in location choice models (figure 3).

The SIMBAD model

This PhD thesis is elaborated within the framework of a LUTI model developed in the LAET, the SIMBAD model (figure 4). The objective of the SIMBAD model is to integrate the sustainable development concerns in the long-term evaluation process of transport and local land-use policies for the urban area of Lyon (Nicolas, 2010). The SIMBAD model is based on the URBANSIM platform (Waddell, 1998) for the land-use sub-model, whereas the transport sub-model is external, developed by the LAET.

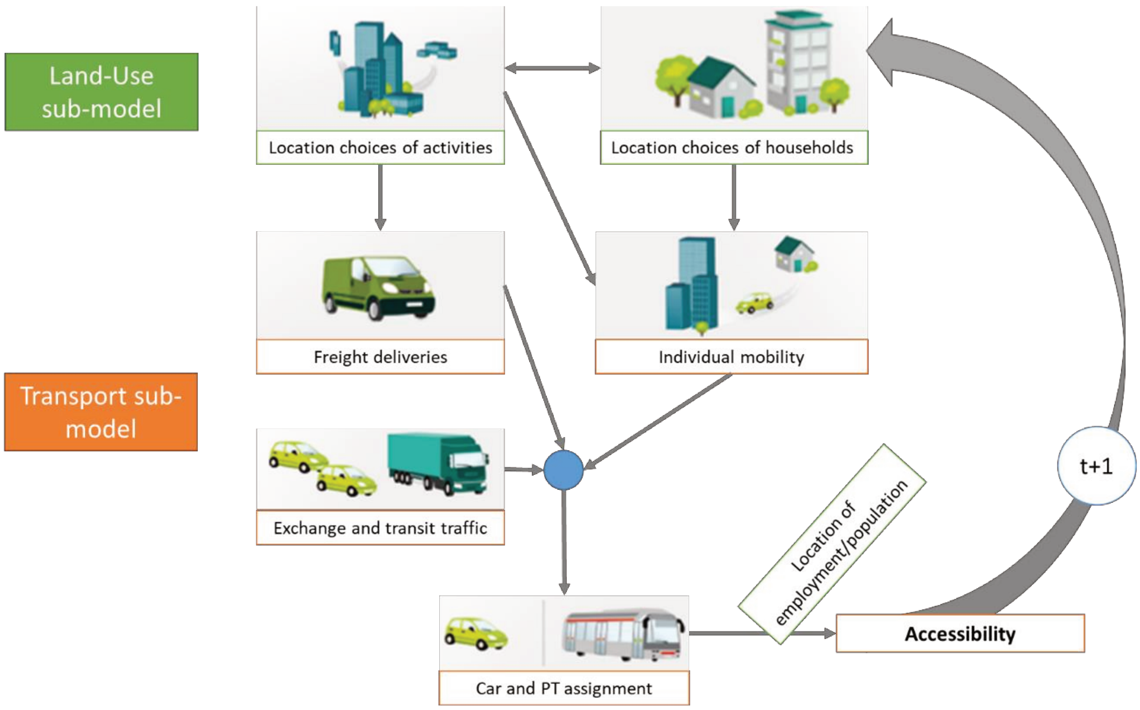


Figure 4 *The structure of the SIMBAD model of the land-use and transport sub-models, adapted from Nicolas, 2010*

The socio-demographic sub-model simulates different socio-demographic characteristics and preferences that change over the years. The outputs of this model feed the land-use sub-model, in order to take into account the updated socio-demographic structure of the population based on the observed tendencies. In the application of the SIMBAD model, those projections are based on the national projections estimated by the INSEE.

The land-use model reproduces different interactions that occur in the urban environment from a market perspective. Central components of the land-use sub-model are the location choice models of households and firms. All the other components of this sub-model provide information on the location choice models (land value, new floor space, etc.) in order to simulate the evolution of the land use based on the choices of households and firms. In the SIMBAD model, the location choice models are based on the RUM theory. Their implementation is based on discrete choice models, which determine the location choice of households and firms. For both agents the logic is the same. As a first step, a model (binomial logit) estimates the probability for each agent to relocate, based on the attributes of the current location (accessibility included) and the characteristics of the agent. As a second step, a location choice model estimates the probability of the different possible locations to be chosen by new and relocating agents, based on the location attributes (accessibility included) of the alternative zones and the characteristics of the agents. The estimated probability serves for the attribution of new locations for new and relocating agents.

Lastly, the transport model uses the socio-demographic structure of the population and the spatial distribution of land use as input, provided by the land-use sub-model. For the urban area of Lyon, the transport model is a traditional four-step trip-based model. The four steps of the modelling procedure are as follows:

(i) the generation of the inbound and outbound flows for each zone;
(ii) the distribution of the flows between origin and destination zones;
(iii) the modal choice for each estimated flow;
(iv) the assignment of the flows on the transport network to determine the generated traffic for each part of the network.

The transport sub-model takes into account the exchange and transit flows between outside locations. Additionally, SIMBAD integrates the flows of freight deliveries, supplied by the FRETURB model (Nicolas, 2010). As an output, the transport model provides generalised costs between origins and destinations. These generalised costs serve the purpose of calculating the appropriate accessibility measures (figure 4).

The key role of accessibility in the modelling chain

Based on the general structure of LUTI models and the SIMBAD model in particular, the generalised costs calculated by the transport sub-model should be integrated into the land-use sub-model to account for changes occurred on the transport system. This integration should affect the attractiveness of locations (based on the feedback loop, figure 2). Generalised costs are meaningless for the location choice models alone. If we consider again that travel is an induced activity, people move in order to reach activities at destinations. The cost itself does not integrate the attraction component of the activities. The accessibility permits the combination of the generalised costs with the number of activities and provides a quantifiable amount of attractiveness of an area.

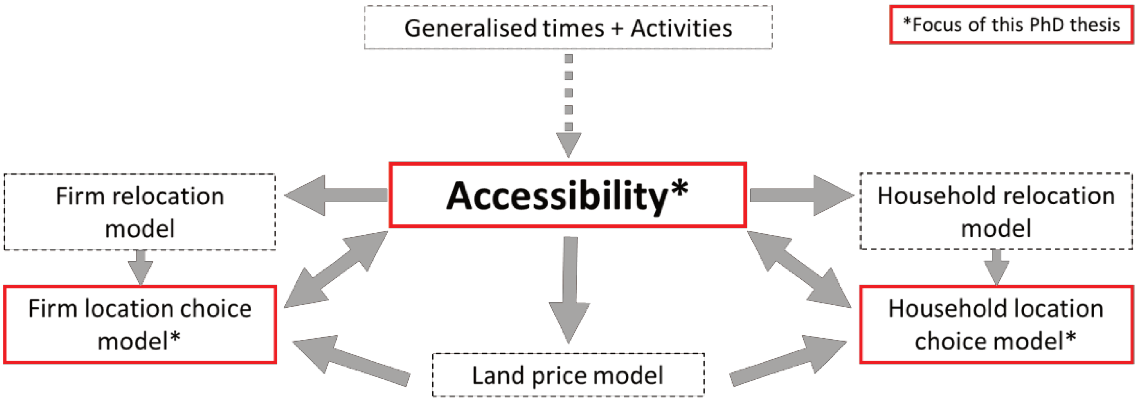


Figure 5 *The key role of accessibility in LUTI models*

Accessibility is then integrated into relocation and location choice models of households and firms (figure 5). A relocation model identifies the potential relocating agents. Then, a location choice model locates new and relocating households and firms. Between other location factors, households are sensitive to accessibility to employment (Alonso, 1964; Baraklianos et al., 2018b; Lee et al., 2010). Accessibility to employment represents the ease of access not only to working opportunities but also to different activities, such as services, entertainment or shopping. Between other location factors, firms are sensitive to accessibility to population (Baraklianos et al., 2018a; de Bok and Van Oort, 2011; Jacobs, 1969). Accessibility to employment represents the ease of access to potential clients, workforce, suppliers and distributors (Baraklianos et al., 2018a).

The aforementioned description of the modelling process shows that accessibility is ubiquitous and has a major role in the modelling chain of a LUTI model. It permits the consistency between the transport model and the land-use model. It also guarantees the interaction between the locations of households and the locations of firms. In that way, if there is an accessibility change, households and firms will evaluate the possibility to relocate and possibly adapt their locations. An increase of accessibility will permit households and firms to reach more activities or population respectively, at the same or a lower generalised cost (Geurs and Ritsema van Eck, 2001; Mulley, 2014).

Integrating accessibility in location choice models

Accessibility measurements

In relation to the three accessibility components, in order to model accessibility, we need three construction elements: an activity to which accessibility is computed (e.g. employment), a mode by which individuals reach opportunities (e.g. car) and a concerned group for which accessibility is computed (e.g. households without a car) (Bouzouina et al., 2014).

There is a wide selection of different methods to integrate those three elements. The available methods can be grouped into three large families: the **infrastructure-based**, the **opportunity-based**, and the **utility-based** measures (Geurs and van Wee, 2004). Each group focuses more on one accessibility component. The spatial and temporal dimension concern all types of accessibility measures. The spatial character of accessibility is embedded therein since we refer to the accessibility of a location. The temporal dimension can be integrated in all three measurement methods but with different detail level.

Infrastructure-based accessibility measures focus on the transport system supply. They represent the facility to travel using the existing infrastructure. Some examples of infrastructure-based measures are:

- Congestion level;
- Length of motorways;
- Number of passenger-seats per hour;
- Number of metro/tramway stations;
- Distance to a metro/tramway station/motorway;
- Etc.

If we give them a spatial dimension, for example number of passenger-seats per hour in the city centre/outskirts, they approach the accessibility concept more. Nevertheless, despite their wide application, these measures perform poorly in characterising accessibility. For example, a motorway provides great mobility speeds but, in terms of accessibility, we do not have any information. This is because infrastructure-based measures do not integrate the most important component of accessibility, land use. They can measure the performance of a network in terms of mobility but there is no connection with the activities to and from which people travel.

Opportunity-based or location-based measures are the most commonly used accessibility measures for the evaluation of land-use and transport policies. As their name reveals, they highlight the land-use component of accessibility. Due to their aggregate nature, they are relatively easy to apply and to interpret. Based on some assumptions, they can integrate all the accessibility components and dimensions. They can even account for capacity constraints (Bunel and Tovar, 2014). The two most frequently applied opportunity-based measures are the cumulative opportunities and the potential or gravity accessibility measure. The cumulative opportunities measure can be considered a special case of the potential accessibility measure (Handy and Niemeier, 1997). In general, these indicators measure the number of opportunities that can be potentially reached from an origin, given the transport options, based on an impedance function. A reduction of cost due to a project can increase accessibility. Under the hypothesis that the satisfaction of individuals increases with the availability of opportunities, potential accessibility measures are used for project evaluation. They can measure the consumer surplus variation resulting from a new transport project (Geurs and Ritsema van Eck, 2001). Methods from space-time geography (Cascetta et al., 2016; Kwan, 1998; Miller, 1999; Wang et al., 2018) show that it is possible to integrate temporal constraints to the opportunity-based accessibility measures. One of the critics is the non-integration of individual variability. This group of measures, due to their aggregate nature, captures only systematic taste variations of socio-demographic groups. The underlying hypothesis is that individuals of the same socio-demographic groups have the same behaviour, something that can be misleading.

Utility-based measures bypass the limits of the opportunity-based measures and highlight the individual component of accessibility. Their disaggregate nature and the fact that they derive from a solid theoretical framework, the RUM theory, make them very attractive. In the framework of the RUM theory, accessibility equals the expected maximum utility that an individual derives from an available choice set (Ben-Akiva and Lerman, 1985). Because utility is a random variable, we are referring to the average of the maximum utility or benefit. The basic assumption is that an individual allocates a utility value to all available alternatives as destination choices based on the attributes of the alternatives, the mobility options and the characteristics of the individual. The sum of the utility values is the accessibility that an individual enjoys at a specific location. In practice, one can estimate this “logsum” value using a destination choice or a mode-destination choice model. It is the denominator of the mode/destination choice of a trip-based or an activity-based transport model (Bhat et al., 2000; Guevara and Ben-Akiva, 2006). One can also integrate temporal variables that influence the derived utility (Cascetta et al., 2016). However, despite the many advantages of the utility-based measures, they are not widely applied in empirical studies. The need for detailed individual data and destination choice models makes their application burdensome and costly (Geurs and van Wee, 2004).

Identified challenges

Finding the appropriate accessibility measure is a challenging process. Before modelling accessibility, one needs to make certain simplifications, methodological choices and find the best balance between an ideal measure and possible constraints (figure 6). Ideally, an accessibility measure should be theoretically sound, meaning to have a behavioural basis and to include all the components and dimensions of accessibility in the most exhaustive way. In addition, the measure needs to be consistent with the objectives of the model/study. If the objective is to evaluate new transport infrastructure projects, the accessibility indicator needs to capture the possible variations of the travel cost. If the objective is more socially oriented, the integration of the accessibility needs to reflect those social aspects of the model (e.g. capturing the systematic taste variation of accessibility). If the objective is to compare different levels of sensitivity to accessibility between different groups, the accessibility indicator needs to be the same between the different groups. However, those choices are strongly conditioned by the availability of data, the estimation capacity, the desired interpretability/traceability of the results and the consistency with the transport model (figure 6). In the case of data absence, one needs to make the right choices and hypotheses or find appropriate proxies.

One important constraint that influences methodological choices is the capacity to estimate different accessibility indicators. It is possible that we cannot estimate the most appropriate, from a theoretical point of view, accessibility indicator, which corresponds the best to the objectives of the study, due to data and modelling constraints. A good illustration of this "mismatch" between what is theoretically the best and what is applicable is the use of utility-based measures. Utility-based measures are considered as the most appropriate accessibility indicators for the location choice models because of their theoretical bases and their consistency with the RUM theory. Additionally, they can account for temporal constraints (Cascetta et al., 2016). Their superiority and their ability to capture individual variability has been demonstrated by Dong et al. (2006). However, the utility-based accessibility measure is derived by a destination (or mode-destination) choice model of a transport model. Additionally, the best utility-based measure is estimated by state-of-the-art activity-based transport models, which are not largely available.

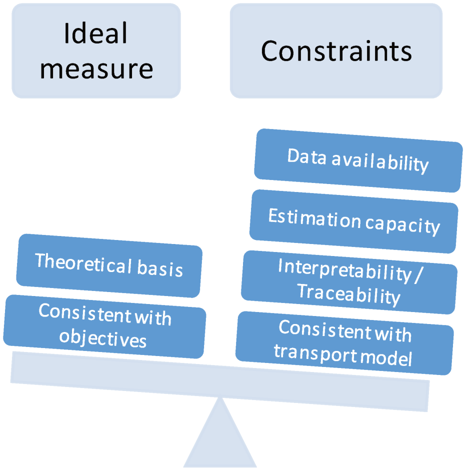


Figure 6 *The balance of an accessibility measure choice*

Schirmer et al. (2014) demonstrated the plurality of accessibility measures for the implementations in residential location choice models. In the firm location choice models, we observe the same plurality. One can either introduce a comprehensive and complete indicator of accessibility, which covers all accessibility components, or introduce various simpler accessibility measures, each covering one or more accessibility components. Within the plurality of approaches and constraints, we do not have a clear picture of the effect of methodological choices on the results of location choice models. While there have been works on the effect of the accessibility measure on policy evaluation (El-Geneidy et al., 2011; Handy and Niemeier, 1997) or other modelling frameworks (Boisjoly and El-Geneidy, 2016; Bunel and Tovar, 2014), there are few works in the literature studying the influence of the accessibility measure on the results of a location choice model. There is a debate in relevant literature concerning the real effect of accessibility in location choice, mostly residential choice (Eliasson, 2010; Lee et al., 2010; Weisbrod et al., 1980; Zondag and Pieters, 2005). As is the case for policy evaluation, the measurement method can affect the modelling and simulation results of a location choice model.

Another important aspect regarding the inclusion of accessibility measures in location choice models is the integration of the individual component (Wee, 2015). Different households and firms value accessibility differently. In addition, as accessibility is a multicomponent measure, different accessibility components can have different influence. A family household could value accessibility differently than a single-worker household. A brand-new firm in the business sector could value differently proximity to a motorway than a relocating logistics company.

New technologies transform how firms organise their functions spatially (Anas et al., 1998). Intuitively, those differences exist but, in the framework of location choice models, are those differences significant? Do we need to develop distinct models or do we need to use different accessibilities for each socio-economic group?

Lastly, a dimension not thoroughly analysed in the literature of the location choice is the temporal evolution of the preferences. As new generations become active citizens in different economic and technological contexts than the previous ones, it is possible that their preferences evolve. Literature indicates that newer generations showcase decreasing car ownership and use and an increasing preference for public transport (Wee, 2015). Such changes that have already been studied in the framework of individual mobility have great implications on the location choice preferences. There are already some indications that the young prefer central areas (Melia et al., 2018). Analysing the temporal evolution of the location choice preferences and especially the temporal evolution of the preferences for accessibility can give some important guidelines for anticipating the future of urban environment.

Four essays on accessibility in location choice models

The aforementioned challenges consist important research issues in the context of LUTI and location choice models. The identified challenges can be grouped in three research axes:

(i) How the inclusion of different components in the measurement method influences the location choice modelling results?

(ii) What are the implications of individual taste variation on the location choice modelling?

(iii) What is the temporal stability of location choice preferences?

These three research axes consist the backbone of the thesis. Having them as a springboard, the analysis gave birth to four independent research articles. Each one of them can be related to more than one research axes (figure 7), but all the articles concern issues related to accessibility in location choice models. Two of them use the residential location choices as case studies and two have been based on firm location choices. Nevertheless, the choice of focus on households or firms could be interchangeable as the identified challenges around accessibility concern both households and firms. The research questions of the articles are:

(i) **Chapter I** (Paper 1): Does the accessibility measure influence the results of the residential location choice modelling?

(ii) **Chapter II** (Paper 2): Do new and relocating firms have different preference for accessibility?

(iii) **Chapter III** (Paper 3): What is the impact of accessibility on the location choices of business services?

(iv) **Chapter IV** (Paper 4): Renters vs owners. How has the preference for accessibility evolved for a residential location choice?

Research axes	**Accessibility measurement**	**Individual taste variations**	**Temporal evolution**
Paper 1	●	●	
Paper 2	○	○	
Paper 3	○	○	
Paper 4		●	●

● **Residential location**

○ **Firm location**

Figure 7 *Relation of the PhD thesis papers to the research axes*

Does the accessibility measure influence the results of residential location choice modelling?

In the first chapter of the PhD thesis, the objective is to analyse if the accessibility measure affects the results of residential location choice models. Accessibility is essential in land-use transport interaction frameworks. For residential location choices in particular, it has always been important at the theoretical level. At the empirical level, the place of accessibility has been questioned in some works, considering other more important location factors as more important, like the social environment and neighbourhood amenities. However, this result can be caused by the measurement of accessibility. In view of the wealth of approaches, this paper examines whether different accessibility measures can lead to divergent results. Using a residential location choice model for the Lyon urban area in France, we tested various accessibility indicators and we compared the results. We concluded that accessibility is an indispensable variable. Without it, the model gives inconsistent results. Complex accessibility measures give better results but simple measures are also relevant for residential location choices modelling. The choice highly depends on the objectives of the application, especially if the model is to be used for simulation.

Do new and relocating firms have different preference for accessibility?

The second chapter of the PhD thesis aims to analyse any differences of accessibility preferences between newly created and relocated firms. Accessibility is one of the most important factors for the location choice of a firm. However, even if it seems intuitive, works analysing any differences between firm creations and relocations are scarce. This paper examines whether any important differences exist in an intraurban setting, the Lyon urban area. We rely on discrete choice models and we use data from more than 43.000 creations and 11.000 relocations, having taken place during the period 2005-2011, from eight economic sectors. The results demonstrate that the effect of accessibility differs between firm creations and relocations of the same economic sector. This difference depends on the type of economic activity of the sector and the type of accessibility.

What is the impact of accessibility on the location choices of business services?

The aim of this paper is to evaluate the impact of accessibility on the firms' location choices in the business services sector. Distinguishing between Front Office and Back Office services, we estimate multinomial logit models based on the data of Lyon. The results show that the effect of accessibility differs between economic subsectors. In general, Front Office services prefer highly accessible locations with good transport infrastructure where location externalities are strong because of the importance of face-to-face interactions. Back Office services are sensitive only to the proximity to motorways. In the case of relocations, all establishments tend to relocate near their previous location.

Renters vs owners. How has the preference for accessibility evolved in residential location choices?

This paper analyses the temporal evolution of the preferences for accessibility. Urban areas face important challenges. More and more people choose to buy a residence in the suburbs taking advantage of the accessibility increase. At the same time, young households choose to rent in central areas. In this paper, distinguishing between renters and owners, we investigate the evolution of the households' location choice preferences over time with a special focus on accessibility, for the urban area of Lyon. We rely on discrete choice models using desegregated census data of the location choices of households from 1999, 2008 and 2013 and we calculate elasticities. The results confirm our initial intuition. Owners become less sensitive to accessibility over time while renters show the opposite. Our results suggest that model scientists and planners incorporate these temporal evolutions into their analyses for a better land-use transport integration and policy.

Transversal choices

The four papers represent the main work of the thesis. Each article is stand-alone, meaning that each one is based on and presents a specific research question, theoretical background, data and methodological choices. However, some methodological choices are shared between the articles. Those common choices concern the definition of accessibility, the study area and the applied modelling method.

Accessibility is the centre of interest of this thesis. This work considers accessibility to be a location attribute, given the strong spatial perspective of the analysis. **Accessibility is defined** as the utility of households and firms derived from the facility to reach activities or population from a location given the mobility options. This means that the characteristics of households or firms do not influence the enjoyed accessibility levels of a location. However, these characteristics define the influence of accessibility during a location choice.

In order to respond to this strong spatial dimension, the accessibility analysis is based on infrastructure-based, and notably on opportunity-based, measures. At a theoretical level, utility-based measures should be more appropriate for integration into the location choice models. They are consistent with the RUM theory and they capture the individual taste variation better (Ben-akiva and Bowman, 1998). However, such a measure is based on the observed individual mobility and reduces the spatial dimension of accessibility. This means that accessibility is conditioned by choices such as owning a car, which does not constitute spatial information. Additionally, the calculation of a consistent utility-based measure was not feasible from a technical point of view. This work uses generalised times provided by a four-step trip-based transport model, which does not integrate a desegregated mode/destination choice model. While it could be possible to develop such a model, the available data for the study area is not rich enough to develop a destination choice model at the same spatial detail as the location choice models. Furthermore, the development of such a model is beyond the scope of this PhD thesis.

The most comprehensive opportunity-based accessibility measure is potential accessibility (Baraklianos et al., 2018b; Geurs and van Wee, 2004). A challenge with this measure is to combine different transport modes serving an area in one accessibility indicator, in order to avoid any problems of multicollinearity in the location choice models. The main transport modes in the study area are the car and the public transport. To integrate both modes, an aggregation of the generalised times is performed. The accessibility indicator is based on composite generalised times, which is a superior method than single-mode accessibility (for more details see Baraklianos et al. 2018b; Bhat et al. 2000; Bhat et al. 1999). This methodology has been applied in all the articles, modifying the indicator based on whether it concerns the households or the firms.

The analysis is performed on the same **study area** in all four papers. All the developed models and analyses concern the urban area of Lyon (figure I.1, page 44). The urban area of Lyon is the second largest urban area of France in terms of both population and economic activity and constitutes an interesting case study. It is a dynamic area, which concentrates an important number of transport investments and innovative land-use policies. The area offers a rich social and economic environment that evolves over time. The choice of the study area was also made for consistency reasons. The works of this thesis are part of the SIMBAD LUTI model, which has been developed for the urban area of Lyon. The ultimate objective is to integrate the contributions of this thesis into the modelling chain of the SIMBAD model.

The **location choice model** applied in all research articles, is a MultiNomial Logit (MNL) model. The retained alternatives are neighbourhoods or, as they are defined by the INSEE, "*grands quartiers*". There were 432 zones in 1999 and 2006, and 431 zones in 2013. While this zoning is not the most detailed for which data is available, it is the most appropriate for the estimation of a location choice model because it retains a certain socio-economic homogeneity (more details in Baraklianos et al. 2018b: Chapter I, page 44).

While the modelling method is the same in all the applications, some details concerning the estimation of the models have evolved throughout the articles. The MNL model in Paper 1 has been estimated using a random sample of alternatives. The high number of alternatives combined with a high number of observations made the estimation without sampling very cumbersome. However, this limitation is overcome in Papers 2 and 3, where the smaller number of observations for firms and the better computing power made an estimation without random sampling of alternatives possible. The last evolution, in Paper 4, was the inclusion of spatially lagged parameters for some variables. The objective was to capture the effect of the spatial autocorrelation of certain variables in the deterministic part of utility, which otherwise could give biased parameters.

Chapter I
Does the accessibility measure influence the results of residential location choice modelling?

1. Introduction

Accessibility is central to Land-Use Transport Interaction (LUTI) models (Acheampong and Silva, 2015; Bonnel et al., 2013; Zondag et al., 2015), playing an important double role. On the one hand, it is one of the main results of the simulation process, facilitating decision-making. On the other hand, it is one of the key variables in the location choice models of households and firms. Accessibility expresses the main effect of the transport system (Zondag et al., 2015) quantifying the potential interaction between land use and transport (Hansen, 1959).

At the theoretical level, accessibility is a key determinant in residential location choice models (Alonso, 1964; Lowry, 1964). In empirical models however, a significant relation between accessibility and residential location choice is rather hard to be proven (Blijie, 2005; Lee et al., 2010). Some studies have questioned the importance of accessibility concluding that other location factors like social environment, neighbourhood amenities and dwelling characteristics are more important (Blijie, 2005; Chen et al., 2008; Sener et al., 2011; Zondag and Pieters, 2005). Others, consider accessibility as essential in the estimation of residential location choice models (Eliasson, 2010; Srour et al., 2002) and conclude that accessibility is important even for a polycentric urban structure (Lee et al., 2010). Many factors can contribute to these divergent results like modelling choices (analysis level, explanatory variables, model structure and market segmentation) or local particularities (in areas where the transport services are good, the importance of accessibility tends to decrease). These controversial results might be explained by the fact that the definition and measurement of accessibility has not been thoroughly examined in LUTI and location choice literature.

In the context of the residential location choice, a great variety of accessibility measures is applied (Schirmer et al., 2014), from simple to complex ones. A simple measure can be the proximity to transport infrastructure or a Euclidian distance of a location to city centre, as it was implemented in the first works of Alonso (1964). Recent modelling techniques like activity-based transport models permit the integration of individual constraints, preferences and ability to travel into accessibility indicators. However, "*accessibility is often a misunderstood, poorly defined and poorly measured construct*" (Geurs and van Wee, 2004, p. 127). Translating its influence in a residential location choice context, where the decision depends on various dimensions, can be a complex task. The measurement method of accessibility can potentially influence the conclusions that one draws on the importance of accessibility.

Recent studies have questioned the sensitivity of the modelling results to the accessibility measure in other frameworks. Boisjoly and El-Geneidy (2016), using public transport share regression models, analysed the influence of time sensitive accessibility measures to the results of modal choice modelling. Bunel and Tovar (2014), in the framework of spatial mismatch, examined if the results of local job accessibility modelling depend on the measurement strategy. But in the context of residential location choice modelling, to our knowledge, little research has analysed the effect of the accessibility measure. Guo and Bhat (2007) studied the impact of the definition of the alternative zones on the results of a residential location choice model and accessibility was a variable of interest. Srour et al. (2002) for Dallas-Fort Worth applied two different accessibility measures (cumulative opportunities and logsum from a trip-based model) to three activities (work, green space, shopping).

Our objective is to extend Srour et al. (2002) work, analysing the influence of accessibility measurements on the results of a residential location choice model. More precisely, first we want to analyse the importance of accessibility, even when one applies simple measures and second, to assess the benefit of using more sophisticated measures. We developed an empirical application for the urban area of Lyon. Our application is based on previous works which allowed to develop a residential location choice framework (Aissaoui, 2016; Aissaoui et al., 2015; Kryvobokov and Bouzouina, 2014).

Usually, the decision on the best accessibility measure is based on statistical indicators of the model. In our work, we take a step forward and we also analyse the market shares predictions of location choices. When developing models for planning, it is important to get good statistical indicators but also to reproduce correctly the observed market share. Our work aims to provide guidance to modellers and decision-makers on making better and faster decisions regarding the use of accessibility in residential location choice modelling.

The rest of the paper is structured as follows. Section 2 presents the accessibility theory and the applied measurement methods in the framework of residential location choice studies. Section 3, “Methodology and application”, summarises our methodological choices and presents the study area and the data used for our empirical work. Section 4 analyses the modelling results while section 5 summarises the results and discusses the conclusions of the article.

2. Accessibility and residential location choice: one concept, different approaches, different results?

The concept of accessibility is difficult and complex due to the fact that it is a multicomponent and a multidimensional construct (Boisjoly and El-Geneidy, 2016; Cascetta et al., 2016; Geurs and van Wee, 2004; Niedzielski and Boschmann, 2014) for which the measurement methods are not yet standardised (Acheampong and Silva, 2015).

Accessibility is a construct of three components, transport, land-use and individuals (Niedzielski and Boschmann 2014). These three components evolve in two dimensions, in space and time (Geurs and van Wee, 2004). Locations that increase their accessibility are able to attract more activities and population to move in (Axhausen, 2008). An analysis of the accessibility changes for over 150 years in Switzerland have shown how transport networks, population and activities evolve together over time (Axhausen et al., 2011).

Initially, accessibility was not conceived as a multicomponent and a multidimensional concept. It was merely a two-component concept within the spatial dimension (Hansen, 1959), defined as the interaction between transport and land use. Further research highlighted the importance of individual component (Ben-Akiva and Lerman, 1985) and temporal dimension (Hägerstrand, 1970). Accessibility gains on theoretical developments were posterior to advances of empirical methods in several fields, particularly in economics and geography. This led to a fragmentation of the applied methods, points of view, methodological processes and assumptions[1].

[1] For a review on the contrasts of the accessibility approaches see Niedzielski and Boschmann (2014)

Translating these components and dimensions of accessibility into indicators is not an easy task. For a residential choice, transport infrastructure could be an opportunity for household members as it increases their ability to travel. However, it can create negative externalities such as pollution, which can discourage a household from locating in close proximity or can incline people to move out (De Palma et al., 2007; Hamersma et al., 2015). A residential location is also affected by the spatial distribution of the activities in which household members participate like employment, shopping and leisure (Wegener and Fürst, 1999). Households' members in areas with plethora of activities can make shorter and more optimised travels due to trip-chaining (Hu, 2017) and have more time to participate in leisure activities (Cordera et al., 2017). Individual preferences and abilities influence activity participation and transport choices. All these choices are constrained by the individual space-time activity prism (Hägerstrand, 1970), which limits the ability of individuals to participate in all desired activities within the limited time of the day. Accessibility indicators that incorporate individual component and temporal dimension, require data and modelling techniques that are not always available.

In residential location choice literature, the transport and land use components are integrated through location-based accessibility measures (distance to centre, cumulative opportunities, potential or gravity based accessibility), which mostly capture the spatial dimension of accessibility. The individual component is integrated either by interacting location-based accessibility measures with households' characteristics to capture systematic taste variations or by using logsum accessibilities, derived by trip-based or activity-based transport models[2]. The temporal dimension can be integrated by space-time accessibility measures that calculate the available opportunities given the time constraints of the individuals. However, due to limits of available spatio-temporal data, the temporal dimension is usually integrated within the transport component (peak-on times). Newly available geolocation-based data from various sources (smartphones, smartcards etc.) provide new opportunities in analysing the temporal dimension of accessibility in a dynamic and comprehensive manner (García-Albertos et al., 2018; Tenkanen, 2017). In Table I.1, we present some empirical studies of residential location choice models to illustrate the diversity of the measures and the conclusions on the importance of accessibility. We have not set out to be exhaustive but to give some representative recent works, which illustrate our research question.

[2] For a general review of accessibility measures see Geurs and van Wee (2004) and for a review of the accessibility measures applied in residential location choice models see Schirmer et al. (2014)

In Paris, De Palma et al. (2007), found that the distance to centre, the distance to motorways and the number of railway and subway stations were significant. The authors highlighted that the accessibility to transport infrastructures is more important than the negative externalities they can cause. In Dallas-Fort Worth, Srour et al. (2002) using logsum accessibilities and cumulative opportunities measures, found that accessibility to employment is more important than accessibility to shopping and green space and concluded that cumulative opportunities measure is the most appropriate. For the same study area, Guo and Bhat (2007) found that the perception of accessibility can vary based on the households' characteristics. In general, accessibility to employment has a negative impact on a location choice, but higher income and one-individual households tend to choose locations with good accessibility to employment. This outcome can be the effect of the housing price. Areas with high accessibility tend to be more expensive (Coppola and Nuzzolo, 2011) and thus more attractive to rich households. In Stockholm, Eliasson (2010) using logsums calculated by an activity-based model, found that the attractiveness of a location was positively influenced by the accessibility to workplaces and to other activities like services, shopping etc. The author concluded that accessibility is key in location choice models but one must include various activities and not only employment. In Mecklenburg County, a polycentric urban area, Cho et al. (2008) using a logsum accessibility to ten different employment sub-centres (employment hubs that were identified using spatial econometric methods), found that in general accessibility to employment is a determinant factor for a residential location choice. Then, using market segments models by income, they found that households appreciate accessibility to different employment sub-centres depending on the specialisation of the centre. In their application for the Puget Sound region, Lee et al. (2010) applied together three different accessibility measures (cumulative opportunities for shopping, logsum accessibility to work and space-time prism for shopping). They found that all the accessibility measures matter. More precisely, the logsum for trips to work, estimated by an activity-based transport model, was significant for residential location choice, even after controlling for other location, neighbourhood and dwelling attributes. Both the cumulative accessibility for shopping opportunities and the shopping opportunities within the space-time prism from work to home trips are important. This work, to our knowledge, was the only application that integrated a space-time measure in a residential location choice model. The fact that this type of measure is not commonly applied in residential location choice modelling may be due to empirical difficulties (Boisjoly and El-Geneidy, 2016; Cascetta et al., 2016; Geurs and van Wee, 2004).

Some studies question the importance of accessibility for residential location choices. In the Netherlands, Blijie (2005) found that the distance to motorway ramps was significant for three household types and that the household's car ownership influences this sensitivity. This relation with the car ownership and the residential location choice has been identified in the literature as a self-selection bias; people who like using their car are likely to choose a car-friendly neighbourhood (Cao et al., 2009). At the same time, the distance to a railway station was significant for only one household type. The author argued that accessibility has a marginal influence on residential location choice. Zondag and Pieters (2005), for the same study area and using a similar modelling approach (market segments models), found that activity-based logsums for work and education trips did not have any significant influence. However, the logsums for "all trip purposes" or "other trips" were significant only for some household types. Their conclusion was that accessibility has a minor influence on residential location choice. In their application for the San Francisco Bay Area, Sener et al. (2011) applied three different accessibility measures (infrastructure based, individual accessibility and a cumulative access to work measure) and used two different modelling strategies (Multinomial logit and Distance-based Spatially Correlated logit). They found that the zonal motorway density was significant only when the Multinomial logit model was used, while the household commute time to work (the sum of working household members) had a significant negative influence. Moreover, the number of household members with work location within 30 minutes by Public Transport was positive and significant. However, they mentioned that location-based accessibility measures were insignificant, without specifying the applied accessibility measure. Another study for the Dallas-Fort Worth metroplex (Guo, 2004), using potential accessibility measures to employment, shopping and leisure opportunities, found that only accessibility to shopping opportunities was significant for all households while the accessibility to the other two activities were sensitive to household characteristics (education level of the head and the race of the household). They concluded that accessibility to general employment is not important except for the educated workers. For the city of Santander in Spain, Ibeas et al. (2013), using three different model structures (multinomial logit, nested logit and cross-nested logit), found a non-significant parameter of the potential accessibility to employment. They argued that the included variable "individual time to work" was capturing all the effect of the accessibility. In the Chicago Metropolitan Area, Hu and Wang (2017) examined the influence of the accessibility for residential location choices of poor households. Using a multinomial logit model and a potential accessibility to employment, they found that in general, car or Public Transport accessibility was not significant for location choices. Only after interacting accessibilities with the car ownership, they found a negative influence of the accessibility by car for the car owners and a positive influence of the accessibility by Public Transport for the households without a car. The first counterintuitive result is possibly related to the effect of the land value, since poor households are very sensitive to this location factor. However, there can be some endogeneity issues concerning the car ownership and the preference for accessibility due to self-selection bias discussed earlier.

The results in the literature are diverging. There is a need to study the impact of the accessibility measure on the results of residential location choice modelling *ceteris paribus*. The existing studies have applied many different accessibility indicators to different contexts, so their results are not comparable. Potentially, the measurement method can influence this result. Therefore, in this paper, we set a threefold objective. Firstly, we test various accessibility indicators to examine if the measurement method influences the results. Secondly, we analyse if different households based on their socio-demographic characteristics appreciate accessibility differently and how this fact affects the modelling results. Thirdly, we examine the ability of the model to replicate the observed choices. This indicator of model quality is important because models need to be validated for their simulation performance.

***Table I.1** Literature review summary*

Source	Study area	Model structure	Multiscale analysis	Market segmentation	Accessibility measures	Interaction accessibility with household characteristics	Conclusion
Blijie, 2005	The Netherlands	Multinomial Logit	No	6 groups	Distance to railway stations and to motorway ramps	Car ownership	Accessibility is significant only for a minority of the estimated models. It has a marginal influence on residential location choices.
De Palma et al., 2007	Ile-de-France (Paris region)	Multinomial Logit	Yes	No	Number of railway and subway stations, distance to motorways, distance to city centre	N/A	Accessibility to transport infrastructure is more important than the local externalities they cause.
Sener et al., 2011	San Francisco Metropolitan Area	Multinomial Logit, Spatially Correlated Logit	No	No	Location-based accessibility, zonal motorway density (km/km²), number of household members with work location in 30 minutes or less by Public Transport	N/A	Accessibility measures were insignificant and important for residential location choices.
Guo and Bhat, 2007	Dallas–Fort Worth metroplex	Multiscale Logit	Yes	No	Potential accessibility to employment, shopping and recreation	Household income, number of household members	Accessibility to employment is important but households have different sensibility based on their socio-demographic characteristics.
Srour et al., 2002	Dallas–Fort Worth metroplex	Multinomial Logit	No	No	Cumulative opportunities and logsum (trip-based) accessibilities to employment, to shopping and to park space	N/A	Accessibility is an important explanatory variable and impacts location choices.
Eliasson, 2010	Stockholm region	Nested Logit	No	No	Logsum (activity-based) for work and other trips	N/A	Accessibility is key and positive location attribute for explaining residential location choices.
Zondag and Pieters, 2005	The Netherlands	Nested Logit	No	6 groups	Logsum (activity-based) to work, education and other trips	N/A	Accessibility has a significant but minor influence on residential location choices.
Lee et al., 2010	Puget Sound Region	Multinomial Logit	No	No	Cumulative opportunities to shopping employment, logsum for work trips (activity-based), shopping employment opportunities in the space-time prism for work to home trips	N/A	Accessibility was significant and important for residential location choice, even after controlling for other location, neighbourhood and dwelling attributes.
Ibeas et al. 2013	Santander, Spain	Multinomial Logit, Nested Logit, Cross-Nested Logit	No	No	Potential accessibility to employment	High income households	Accessibility was not a significant variable for explaining residential location choices.
Hu and Wang 2017	Chicago, USA	Multinomial Logit	No	No	Potential accessibility to employment	Vehicle ownership, working/unemployed	Accessibility to employment in general was not significant. Sensitivity depends on the car ownership status.

3. Methodology and application

3.1. Residential location choice model

The modelling method used in this study is based on discrete choices. In discrete choice modelling, the decision-maker selects the alternative from a choice set which maximises his utility (McFadden, 1977). In residential location choice modelling, the decision-maker is the household, and the alternatives can be large zones, small neighbourhoods or even residential units. The household utility is composed by a deterministic observable part and a random term (equation I.1).

$$U_{in} = V_{in} + \varepsilon_{in} \quad \text{(I.1)}$$

U_{in}: Utility of household n at location i
V_{in}: Deterministic part of the utility
ε_{in}: The error term

Under the assumption that the errors are independently and identically distributed (McFadden, 1977), the probability of a household n making the choice i from a choice set j, takes the logit form as is given by equation I.2.

$$Pn(i) = \frac{e^{V_{in}}}{\sum_{j \in D_n} e^{V_{jn}}} \quad \text{(I.2)}$$

V_{in}: Deterministic part of the utility of household n at i
V_{jn}: Deterministic part of the utility of alternatives j in D_n
D_n: the random choice set

In the Multinomial Logit (MNL) model, the deterministic/observable part of the utility depends on the attributes of the alternatives (zonal, dwelling etc.) and on the socio-demographic characteristics of the households. The utility function takes the form of equation I.3.

$$V_{in} = \alpha X_i + \beta Z_{in} \quad \text{(I.3)}$$

V_{in}: Deterministic part of the utility of household n at i
X_i: A vector of zonal attributes
Z_{in}: Interaction terms of socio-demographic characteristics of household n with the attributes of alternative i
α, β: Parameters to be estimated

The applied residential location choice model in our study is a MNL model and the alternatives are neighbourhoods. The study area is divided in 432 neighbourhoods in order to minimise the spatial autocorrelation effects (see more details for the zoning in section “Study area and data”). The estimation of a model with such a high number of alternatives is computationally difficult. When there is such a large number of alternatives, the parameters of a MNL model can be estimated using a random sample of alternatives D_n of the true choice set C_n and get consistent parameters (McFadden, 1977). We tested for various sample sizes of the D_n, up to fifty choices using various sampling strategies. We concluded that the best sample for the estimation is a random sample of seven random choices, the observed choice included, for every observed household choice (Aissaoui et al., 2015). This small sample is sufficient and gives robust estimations with stabilized parameters.

A limit of the logit model is the assumption that the error terms are Identically and Independently Distributed (IID), which is unlikely in a spatial context (Ibeas et al., 2013). Other modelling structures like nested, cross-nested or mixed logit relax this hypothesis. However, these structures need an a priori assumption on the correlation structure, which possibly does not eliminate completely the problem of spatial autocorrelation (Hu and Wang, 2017), and are difficult from an estimation point of view. Additionally, Ibeas et al. (2013) did not report important differences for the accessibility parameters between the multinomial logit, the nested logit and the cross-nested logit models. Previous works tried to apply a nested logit for our study area, testing for various nesting strategies, but failed to find a better fit than the multinomial logit model (Aissaoui, 2016).

3.2. Accessibility measures and implementation

We consider accessibility as a location attribute so we are relying on location-based accessibility measures. The measures included in this study are the proximity to transport infrastructures, the linear distance to city centre, the generalised time to city centre, the cumulative opportunities and the potential accessibility. The objective is to start by including simple measures and then to introduce more complex and theoretically comprehensive ones in order to analyse the results. More specifically:

- The proximity to transport infrastructures is the simplest accessibility measure that we use. It represents the existence or not of a transport infrastructure in the vicinity. This indicator takes into account only the transport component of accessibility.
- The linear distance to city centre introduces a land-use component and a spatial dimension to accessibility. The land-use component is introduced into its simplest form; all opportunities are located into the city centre (Alonso 1964). The spatial dimension provides us with the first global accessibility measure, as the rest of the applied measures, because we have a relative value for all locations. However, this measure incorporates some strong assumptions on the land-use and transport component. It supposes that all opportunities are located in the city centre and that the transport infrastructure is homogenously available in space, which is not realistic.

- The access time to city centre relaxes the assumption on the availability of the transport infrastructure. It accommodates the real transport supply for trips to city centre and it introduces the time dimension of accessibility because times can refer to peak or off-peak periods. Nevertheless, the assumption on the land-use component is still present.
- The cumulative opportunities measure integrates better the land-use component. It is the sum of the number of opportunities within a predefined threshold of time. This means that the transport and the land-use component both contribute on the accessibility value. However, the *a priori* definition of the contour threshold poses a strong assumption on the influence of the spatial distribution of land-use. An opportunity inside the threshold has the same importance whether it is at 2 or 29 minutes from the origin and has no effect when it is outside of the threshold, independently of its size.
- The potential accessibility measure relaxes all the previous assumptions. It represents the sum of the opportunities that can be reached from an origin, weighed by an impedance function. The impedance function can be a function of travel time, so transport and land-use components have a simultaneous effect on accessibility. This measure represents the most comprehensive accessibility measure that we test in this study.

To apply these measures, we need to define how we implement their components, the transport and the land-use components. In other words, we need to define to which activity and for which transport mode we calculate accessibility (Bouzouina et al., 2014).

We restrain our analysis to car and Public Transport (PT) modes. For the proximity to transport infrastructure, we include proximity to motorways and to PT stations as a dummy variable. This variable takes the value 1 if a metro, tramway or railway station is available within a zone or a motorway passes through, otherwise is 0. The linear distance to centre (equation I.4) implicates no choices for the transport component. For the time to centre and the cumulative opportunities measure, we use generalised peak times by car (equation I.5). For the cumulative opportunities measure (equation I.6), the retained threshold is 30 minutes. In “Grand Lyon” (fig. I.1), the mean home-to-work travel time by car in 2006 was 23 minutes (Sytral, 2007). Our study area is larger, so a threshold of 30 minutes is retained, which is the most common threshold in similar studies (Srour et al., 2002; Waddell, 2010). The potential accessibility indicator (equation I.7) is estimated using a negative exponential impedance function with composite generalised times. The negative exponential function is the most appropriate for an urban environment (Geurs and Ritsema van Eck, 2001).

A calculation of an accessibility using composite generalised times allows to take into account more than one transport mode in an area (Geurs and Ritsema van Eck, 2001). For the aggregation of the generalised times, we applied a method developed by Bhat et al. (1999). The result is a composite generalised time, which combines generalised times by car and PT (equation I.8). The generalised time by car is the reference time for all pairs of OD because it is always available. When PT is not available, the generalised time by PT is equal to +∞. The idea is when both car and PT serve an Origin-Destination (OD) pair, the generalised time should be less than the fastest mode. With more mobility options, it is easier to commute between ODs. Thus, accessibility should be higher. This formulation is chosen because it is theoretically consistent with the concept of accessibility as the benefit of using the transport system; the opportunities represent the utility and the time the disutility to reach those opportunities. The increase of transport solutions must be associated with a decrease of the disutility to reach opportunities (Handy and Niemeier, 1997). There exist other methods to calculate composite generalised times, but the chosen method gave the best results in the residential location choice context of Lyon. A limit of the aggregation of car and PT times is that we lose the relative influence of each mode. Possibly accessibility by car and by PT have different effect on location choices. However, the inclusion of both accessibilities in the same model could create collinearity issues.

$$A_i = d_{ij}$$ *d_{ij}: distance in km, with j = city centre* (I.4)

$$A_i = Tcar_{ij}$$ *T_{ij}: generalised time by car, with j = city centre* (I.5)

$$A_i = \sum_j D_j f_{(j)}$$ *D_j: Number of jobs at j; $f_{(j)}$: a function that takes the value 1 when $Tcar_{ij}$ is less than 30 minutes. Otherwise is 0* (I.6)

$$A_i = \sum_j D_j e^{-\beta Tc_{ij}}$$ *D_j: Number of jobs at j; Tc_{ij}: the composite generalised time from origin i to destination j* (I.7)

$$Tc_{ij} = \left(\frac{Tcar_{ij}}{1 + \frac{Tcar_{ij}}{Tpt_{ij}}}\right)$$ *Tc_{ij}: Composite generalised time from i to j; $Tcar_{ij}$: Generalised time for car; Tpt_{ij}: Generalised time for PT* (I.8)

Concerning land-use, we estimate the accessibility indicators to general employment. For the distance and the time to centre, we make the assumption that all employment opportunities are located in the zone of the prefecture of Lyon. The prefecture represents the administrative centre of the city. This zone concentrates various administration services, offices and activities. For the cumulative opportunities and the potential accessibility, we use the number of jobs per zone from the official registry of economic establishments (see next section). We estimate accessibilities only to general employment for two reasons. Firstly, accessibility to employment is important in residential location choice modelling at theoretical level (Alonso, 1964; Lowry, 1964). It influences the prospect of residents finding a job, facilitates activity participation and affects the quality of life (Hu 2017; Niedzielski and Boschmann 2014). Secondly, accessibility to employment is highly correlated with shopping and leisure accessibilities (0.97-0.99) depending on the measure. This correlation is associated to the spatial distribution of the number of jobs (see table I.2). Looking at the global Moran's I indicators in table I.2, we observe that the data is also spatially autocorrelated (the variable by itself) and at a lesser extent spatially correlated (each variable with the others).

Table I.2 *Correlation (spatial correlation - global Moran's I) of the number of jobs per zone*

	Employment	Shopping employment	Leisure employment
Employment	1.00 (0.47)	0.76 (0.33)	0.77 (0.41)
Shopping employment	0.76 (0.33)	1.00 (0.29)	0.79 (0.39)
Leisure employment	0.77 (0.41)	0.79 (0.39)	1.00 (0.49)

A drawback of the location-based accessibility measures is the difficulty to integrate the individual component, which is an important component of accessibility (Geurs and van Wee, 2004). For a residential location, the literature analysis showed that households' characteristics influence the preference for accessibility. In order to capture these systematic taste variations, we interact certain households' characteristics with accessibility measures in the model. The selected characteristics are the employment status of the head of the household and the household size. Other socio-demographic characteristics were tested with less interesting results. The employment status is essential for the impact of accessibility. Students and households with no stable employment are more sensible to accessibility. They do not have a stable job, so high accessibility to employment is essential (Zondag and Pieters, 2005). The number of individuals in a household impacts the effect of accessibility (Guo and Bhat, 2007). The bigger the household is, the higher the need for large dwellings. Usually, areas with high accessibility lack large dwellings.

3.3. Study area and data

Lyon urban area is the second most populated urban area of France after Paris. In total, the urban area had more than 1.7 million inhabitants and more than 750,000 jobs in 2006. The majority of jobs are located in the city of Lyon, more than 40% are concentrated in the area's central municipalities (Lyon-Villeurbanne). Almost 75% are located inside so-called "Grand Lyon", which is made up of the city of Lyon and some suburbs (fig. I.1). The area is divided in 432 zones, the so-called "*grand quartier*" or "large neighbourhood". This zoning is a census breakdown defined by the INSEE (Institut National de la Statistique et des Etudes Economiques - French National Institute for Statistics and Economic Research). These zones have the advantage to be more detailed than the communes in the city centre and to respect a certain amount of homogeneity of the socio-demographic composition of the population (INSEE 2016). This choice reduces the problem of the spatial autocorrelation of the alternatives (Aissaoui, 2016). Their surface varies from around 0.2 km^2 in the city centre to 20 km^2 in the periphery.

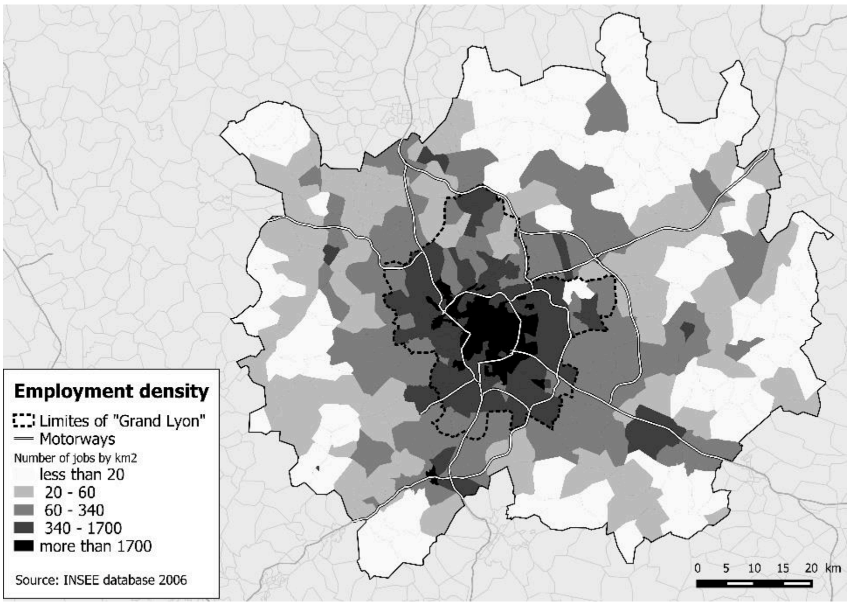


Figure I.1 *Employment density of the study area by zone (classification in quantiles)*

To estimate the residential location choice model, we combine data from different sources. The main data source is the disaggregated census data of 2008, provided by the INSEE, which contains much information about the households like the move-in year, the size, the employment status etc. For the estimation of the parameters, we only use the recently moved households (2006-2008), which provides us with 103,256 observed choices. Table I.3 summarises the distribution of some characteristics of the households introduced in the models.

To characterise the alternatives, we use zonal data provided by the INSEE for the year 2006. We estimate mean housing prices from real estate transactions data (2006) (notary database - Perval).

***Table I.3** Distribution of recently moved households' characteristics*

Household's characteristics	Households	% Households
Quantile of the household's revenue		
rev1 (Households at the 1st quantile of revenue)	26,789	26%
rev2 (Households at the 2nd quantile of revenue)	19,109	19%
rev3 (Households at the 3rd quantile of revenue)	20,931	20%
rev4 (Households at the 4th quantile of revenue)	18,326	18%
rev5 (Households at the 5th quantile of revenue)	17,785	17%
Number of individuals of the household		
1 indiv (Households with 1 individual)	43,625	42%
2 indiv (Households with 2 individuals)	29,003	28%
3+ indiv (Households with 3 or more individuals)	30,312	29%
Employment status of the head of the household		
Act Stable (Households of which the head is active and has a stable activity - permanent contract, public servant)	63,049	61%
Act Not stable (Households of which the head is active but has no stable activity or unemployed - temporary contract, internship, unemployed)	18,328	18%
Not active (Households of which the head is not active - retired, housekeeping, other not active population)	10,405	10%
Student (Households of which the head is student)	11,158	11%

Source: INSEE - General population census 2008

The generalised times by car and PT and the parameter β for the estimation of the potential accessibility indicator (equation 7) were calculated by a transport model developed in the LAET (Laboratoire, Aménagement, Economie, Transport - Transport, Urban Planning, Economics Laboratory) for the Lyon conurbation. The model is a traditional 4-step trip-based transport model with some original enhancements like flows generation with a microsimulation method, three modelling chains based on three revenue groups and the integration of freight flows into the modelling process (Nicolas, 2010). The generalised times and the β value of 0.12 (equation 7) were calibrated using the household travel survey of 2006 (Bouzouina et al., 2014). The employment per zone is calculated using the SIRENE database of 2006, which is the official INSEE company register of all the economic establishments in France. This database is also used for the construction of other variables (see next section). Other accessibility indicators were calculated using Geographic Information Systems using spatial data provided by the IGN (National Geographic Institute of France).

3.4. Model variables and measures

The variables of the model are divided into three categories, the spatial amenities, the social environment and the market trade-off (Aissaoui 2016). The first category concerns the amenities that a household would value to have at close proximity. After various tests, we concluded that the amenities influencing a residential choice in our study area are the presence, or not, of a basic shopping service (bakery, supermarket, convenience store) in the zone, the number of primary schools and the number of secondary schools. We expect that these variables have a positive influence, like in other studies (Chen et al., 2008).

The second category captures self-segregation effects and preference for social housing. Social housing affects the location choices of households. Low revenue households should prefer them but high revenue households might avoid them. The unequal distribution of social housing has played an important role in the residential segregation in France through the concentration of poor households in deprived and stigmatised neighbourhoods (Bouzouina et al., 2018). We introduce this variable as the percentage share of the social housing residential units of the zone (%HLM). Furthermore, the households should have a preference to choose a location that have high concentration of the same revenue (Bouzouina, 2008). This set of variables can reveal a preference for endogenous amenities (Brueckner et al., 1999), which we cannot measure otherwise, like the quality of the services. To measure these self-segregation effects we use the percentage of households in a zone per revenue quantile (%REV3, %REV4, %REV5). We leave out of the analysis the poorest 1st and 2nd quantile of revenue because they are highly correlated with the social housing variables. The data for the construction of the variables come from the databases of the INSEE (see table I.4).

***Table I.4** Descriptive statistics of the zonal attribitues of the variables model*

	Variable description	Variable	Data source	Mean	SD
Spatial amenities	Proximity to basic shopping service (0,1)	Prox Basic Serv	INSEE – SIRENE	0.852	0.356
	Primary schools (number of schools)	Prox Pr. Schools		3.785	3.431
	Secondary schools (number of schools)	Prox Sec. Schools		0.537	0.921
Social environment	Preference for social housing of the 1st quantile of revenue (%)	%HLM*rev1	INSEE – DGI	0.118	0.168
	Preference for social housing of the 2nd quantile of revenue (%)	%HLM*rev2			
	Preference for social housing of the 3rd quantile of revenue (%)	%HLM*rev3			
	Preference for social housing of the 4th quantile of revenue (%)	%HLM*rev4			
	Preference for social housing of the 5th quantile of revenue (%)	%HLM*rev5			
	Self-segregation of the 3rd quantile of revenue (%)	%REV3*rev3	INSEE - Logement	0.245	0.07
	Self-segregation of the 4th quantile of revenue (%)	%REV4*rev4		0.186	0.062
	Self-segregation of the 5th quantile of revenue (%)	%REV5*rev5		0.189	0.121
Market trade-off	Mean zonal housing price (€/m^2)	Housing price	Perval	2,398	269.8
	Accessibility	Acc			
	Proximity to transport infrastructure (0,1)	Transport Infr. Proximity	Authors' calculations		
	Distance to the city centre (km)	Dist. to centre		16.46	11.1
	Time to the city centre by car (minutes)	Time to centre		34.23	17.35
	Cumulative opportunities to employment (number of jobs)	Cumm. Opp.		294,624	241,920
	Potential accessibility to employment (weighted number of jobs)	Pot. Acc.		67,642	69,789

The third category is the market trade-off between accessibility and land value. This category is in the centre of this paper. The observation that households make a trade-off between accessibility and land value has founded the urban economics theory (Alonso, 1964). A household is searching for accessible zones but the land value may discourage a residential choice. We introduce the land value as the zonal mean housing value by square metre. For the accessibility variable, we apply the various aforementioned measures, in order to analyse their impact on the modelling results.

Table I.4 presents some basic descriptive statistics of the variables included into the model. The values concern the zonal attributes.

3.5. Evaluation method

In order to evaluate and compare the results of the models we rely on:

- The likelihood ratio test (equation I.9), using a standard incremental approach in order to analyse the contribution of each variable to the model (Ben-Akiva and Lerman, 1985). The test is applied to nested models, meaning that the unrestricted model contains the same variables of the restricted one plus the accessibility variable to be tested.

$$LRT = -2\ (L(\beta) - L_{\delta}(\beta)) \qquad \begin{array}{l} L(\beta)\text{: Log-likelihood of the restricted model} \\ L_{\delta}(\beta)\text{: Log-likelihood of the unrestricted model} \end{array} \qquad \text{(I.9)}$$

- The comperative analysis of the adjusted rho-squared, which reveals the quality of the model fit to the data.
- The analysis of the relevance of the models' parameters. We analyse if the estimated parameters have the expected signs.
- The ability of the model to replicate the observed market shares. For that, we analyse the differences between the observed market shares and the replicated ones using the Root Mean Squared Error (RMSE - equation 10) (Fox et al., 2014; Washington et al., 2011). The lower the RMSE, the better the performance of the replication. The observed market share is the number of households that chose to move in a specific zone divided by the total number of moved households. The replicated market share is the share of households predicted to move into a specific zone by the model divided by the total number of moved households. To apply the RMSE, we aggregated the initial zoning system from 432 zones to five greater areas (figure I.2). It is more convenient to present the detailed results in five greater areas and to identify the source of the error. The selection of five areas is consistent with the urban structure of the city (centre, periphery) and creates relatevely homogenous zones in terms of socioeconomic composition (Rosales-Montano et al., 2015). The central core of the city, which concentrates most of the economic activity, the east surrounding areas, which are

relatively poor areas but have good transport options (metro/tram), the west surrounding areas, which are considered as rich areas, and the two suburban belts which have lower accessibility levels (see also table I.7 for the mean accessibilities of the different zones).

$$RMSE = \sqrt{\frac{\sum_j (N_j - T_j)^2}{J}} \qquad \begin{array}{l} N_j\text{: Observed market share in } j \\ T_j\text{: Replicated market share in } j \\ J\text{: Number of alternative zones in } C_n \end{array} \qquad \text{(I.10)}$$

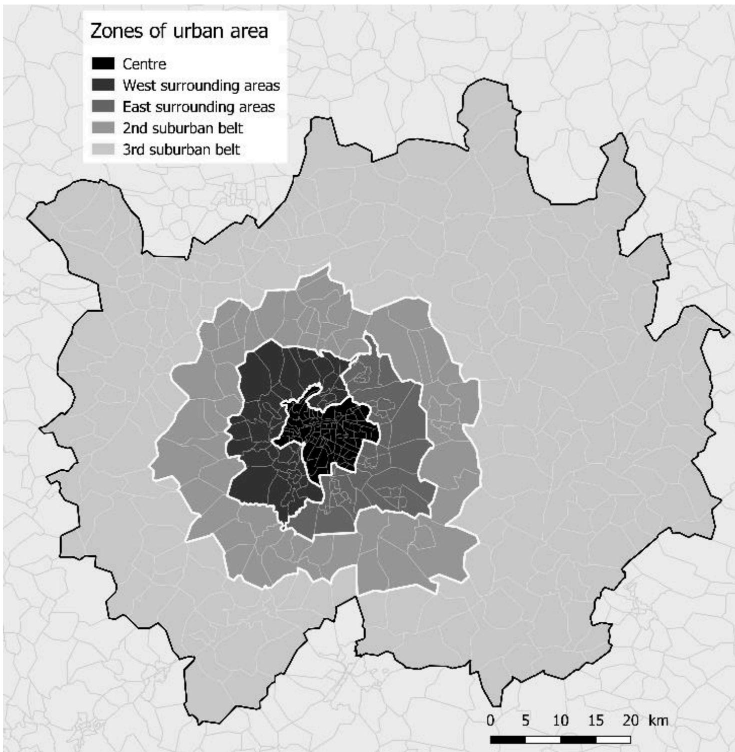


Figure I.2 *The zones of the urban area retained for the RMSE analysis*

4. Results

In this section, we present the results of the estimated models. Tables I.5 and I.6 summarise the results of the models and tables I.8 and I.9 present the model performance to replicate the market shares.

4.1. Analysis of the model parameters

Nearly all parameters are significant in almost all models. Accessibility interacts with the other variables of the model. There are some sign changes and some variables become non-significant.

In the base model, the parameter of the housing price is positive and significant, which is inconsistent with the utility maximisation principle (Guevara, 2015). Housing price should be negative because it represents the disutility of a choice. Otherwise, it would mean that a household would choose the most expensive alternative, everything else equal, which is counterintuitive. Household revenues interacting with social housing from 2^{nd} to 4^{th} quantile are positive, which is not as expected. In model 2, the proximities to metro, tramway and railway stations have a positive effect, which confirms the positive effect of the presence of PT stations, while the proximity to motorways has a negative one. The latter may be due to the negative externalities of motorways such as noise and pollution. These simple accessibility indicators do not correct the non-expected sign of the variables of housing price and social housing proximity. In models 3 to 6, accessibility variables have the expected signs (negative for distance or time to the city centre and positive for cumulative opportunities and the potential accessibility measures). The parameters of housing price are always negative and significant. Household revenues from 2^{nd} to 5^{th} quantile have negative parameters for the proximity to social housing. It seems that the simple definition of accessibility as a proximity to transport infrastructure at local level leads to inconsistent modelling results. When we apply measures that capture the global effect of the accessibility, they give better results regarding the parameters, even simple definitions like the linear distance to the city centre.

Table I.6 presents the results of the models with interaction terms between accessibility and the chosen socio-demographic groups. The objective is to capture some systematic taste variation for accessibility of different households. We present the results for all accessibility measures except of the variables of the transport infrastructure proximity because their inclusion already gave inconsistent results. The different parameter values of accessibility estimators confirm that the preference for accessibility depends on the household’s characteristics. The signs of the estimators are always consistent for all households’ characteristics. Accessibility has a positive influence for a residential location choice for all the selected accessibility measures.

4.2. Model quality and statistical contribution of accessibility

The likelihood ratio test is always significant, meaning that all measures have a statistically significant contribution to the model. Additionally, the integration of interaction terms capturing systematic taste variation has a significant contribution in all models and for all the chosen socio-demographic characteristics.

Table I.5 *Parameters and statistical tests of the models using all accessibility measures*

	Base model (no accessibility)	Transport Infr. Proximity	Dist. to centre	Time to centre	Cumm. Opp.	Pot. Acc.
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Prox Basic Serv	0.93**	0.85**	0.76**	0.81**	0.86**	0.87**
Prox Pr. Schools	0.10**	0.10**	0.11**	0.11**	0.10**	0.10**
Prox Sec. Schools	0.17**	0.15**	0.11**	0.11**	0.13**	0.08**
%HLM*rev1	2.53**	2.14**	0.06	0.08**	0.14**	0.32**
%HLM*rev2	1.75**	1.31**	-0.79**	-0.76**	-0.68**	-0.52**
%HLM*rev3	1.17**	0.69**	-1.37**	-1.31**	-1.25**	-1.06**
%HLM*rev4	1.15**	0.72**	-1.68**	-1.46**	-1.58**	-0.98**
%HLM*rev5	-0.28**	-0.26**	-2.74**	-2.57**	-2.49**	-2.12**
%REV3*rev3	1.72**	1.93**	2.99**	3.15**	2.94**	2.99**
%REV4*rev4	6.06**	7.42**	5.29**	6.35**	5.22**	8.15**
%REV5*rev5	2.26**	3.67**	3.09**	3.30**	2.92**	3.70**
Housing price	1.60**	0.60**	-0.97**	-1.16**	-0.29**	-0.93**
Accessibility	-	-	-0.10**	-0.06**	1.49**	0.54**
Accessibility 2	-	-	-	-	-	-
Motorway	-	-0.17*	-	-	-	-
Metro	-	0.79*	-	-	-	-
Tramway	-	0.69*	-	-	-	-
Railway station	-	0.09*	-	-	-	-
Observations	102,940	102,940	102,940	102,940	102,940	102,940
Log of likelihood zero	-200,312	-200,312	-200,312	-200,312	-200,312	-200,312
Log of likelihood (LL)	-163,899	-154,006	-153,267	-152,671	-155,676	-149,491
Rho-squared	.182	.231	.235	.238	.223	.254
Likelihood ratio test+	Base	vs 1 19,784**	vs 1 21,262**	vs 1 22,455**	vs 1 16,445**	vs 1 28,815**

** significant at 95%

The overall quality of the models increases with the integration of the accessibility components into the indicator. We observe that more complex accessibility measures fit better the data. The potential accessibility indicator has the highest rho-squared, meaning that it captures the best the variation of the households' preferences.

The cumulative opportunities measure (model 5) is generally considered as a better accessibility measure (Geurs and van Wee, 2004) than the distance or time to the city centre because it integrates the land-use component of accessibility, but the rho-squared of the models does not confirm this.

All models with interaction terms have higher rho-squared, meaning that there is a difference in preference for accessibility depending on the households' characteristics. The statistical contribution of the households' characteristics is always significant, based on the likelihood ratio test (table I.6). Between the selected characteristics, the employment status of the household head gives the best fit to the data.

***Table I.6** Parameters and statistical tests of the models including interaction terms*

	Dist. centre * Nb Indiv.	Dist. centre. * HH Status.	Time centre * Nb Indiv.	Time centre. * HH Status.	Cumm. Opp. * Nb Indiv.	Cumm. Opp. * HH Status.	Pot. Acc. Composite * Nb indiv.	Pot. Acc. Composite * HH Status
	Model 3 - 1	Model 3 - 2	Model 4 - 1	Model 4 - 2	Model 5 - 1	Model 5 - 2	Model 6 - 1	Model 6 - 2
Prox Basic Serv	0.79**	0.83**	0.84**	0.88**	0.89**	0.93**	0.89**	0.92**
Prox Pr. Schools	0.11**	0.11**	0.11**	0.11**	0.11**	0.11**	0.10**	0.10**
Prox Sec. Schools	0.11**	0.11**	0.10**	0.11**	0.13**	0.13**	0.08**	0.09**
%HLM *rev1	0.01	0.14**	0.09**	0.27**	0.15**	0.18**	0.39**	0.55**
%HLM *rev2	-0.80**	-0.78**	-0.73**	-0.69**	-0.61**	-0.58**	-0.42**	-0.40**
%HLM *rev3	-1.36**	-1.37**	-1.26**	-1.26**	-1.16**	-1.14**	-0.95**	-0.96**
%HLM *rev4	-1.54**	-1.55**	-1.27**	-1.26**	-1.29**	-1.32**	-0.75**	-0.77**
%HLM *rev5	-2.43**	-2.53**	-2.21**	-2.32**	-2.22**	-2.34**	-1.73**	-1.85**
%REV3*rev3	3.02**	2.71**	3.19**	2.82**	2.82**	2.62**	2.99**	2.65**
%REV4*rev4	5.75**	5.83**	6.86**	6.87**	5.76**	5.81**	8.41**	8.43**
%REV5*rev5	3.46**	3.39**	3.72**	3.61**	3.21**	3.12**	4.00**	3.90**
Housing price	-1.03**	-1.08**	-1.20**	-1.22**	-0.28**	-0.27**	-0.91**	-0.93**
Acc * 1 or 2 indiv	-0.13**		-0.08**		1.88**		0.65**	-
Acc * 3+ Indiv	-0.06**		-0.03**		0.77**		0.28**	-
Acc * Mid Educ							-	-
Acc * Low Educ							-	-
Acc * High Educ							-	-
Acc * Act No Stable		-0.15**		-0.09**		2.22**	-	0.75**
Acc * Act Stable		-0.08**		-0.05**		1.18**	-	0.43**
Acc * No active		-0.09**		-0.06**		1.38**	-	0.49**
Acc * Student		-0.39**		-0.19**		5.56**	-	1.38**
Observations	102,940	102,940	102,940	102,940	102,940	102,940	102,940	102,940
Log of likelihood zero	-200,312	-200,312	-200,312	-200,312	-200,312	-200,312	-200,312	-200,312
Log of likelihood (LL)	-150,704	-148,923	-149,555	-148,083	-153,604	-152,142	-146,507	-145,315
Rho-squared	.248	.257	0.253	0.261	0.233	0.241	.269	.275
Likelihood ratio test^{+}	vs 3 5,126**	vs 3 8,689**	vs 4 6,231**	vs 4 9,175**	vs 5 4,144**	vs 5 7,068**	vs 6 5,968**	vs 6 8,352**

** significant at 95%

4.3. Ability to replicate market shares

In this chapter, we analyse the capability of the model to replicate the observed market shares (Tables I.8 and I.9). We present for each model the difference between observed and replicated market shares and the general RMSE aggregating the 432 zones into five greater zones (fig. I.2). Table I.7 summarises the descriptive statistics for the accessibility variables for each greater area.

Most of the models struggle to replicate the share of the city centre. This is probably because the control variables of the model and some accessibility measures cannot capture the importance of the centre for a residential choice in comparison to other zones. Only when we include the potential accessibility, the error for the central zone decreases significantly. The potential accessibility controls better for systematic prediction errors than the other accessibility measures.

***Table I.7** Descriptive statistics of the accessibility variables for the aggregated zones*

	Transport Infr. Proximity								Dist. to centre		Time to centre		Cumm. Opp. car		Pot. Acc. Composite	
	Motorway		Metro		Tramway		Railway station									
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Centre (Lyon - Villeurbanne)	0.19	0.39	0.67	0.47	0.34	0.48	0.29	0.46	2.71	1.28	11.38	4.44	606,229	36,791	183,057	33,195
East Areas	0.55	0.50	0.07	0.26	0.24	0.43	0.19	0.40	7.54	1.89	22.49	3.82	581,811	53,898	116,327	26,736
West Areas	0.40	0.49	0.02	0.15	0.00	0.00	0.44	0.50	6.36	1.90	19.36	5.36	509,690	109,197	108,639	31,321
2nd Belt	0.34	0.48	0.00	0.00	0.00	0.00	0.34	0.48	13.56	2.06	31.80	4.13	270,529	132,785	44,605	23,998
3rd Belt	0.22	0.41	0.00	0.00	0.00	0.00	0.16	0.36	26.54	6.38	49.21	10.48	78,584	57,005	11,835	6,676

There is a relation between model quality measured by rho-squared and replicated market shares. Model 1 has the highest RMSE meaning that it gives the worst replications. As stated, most of the error comes from the replication of the share of the centre. When we use the transport infrastructure proximity (model 2), the RMSE is relatively high as well. This definition of accessibility cannot capture the utility of the city centre and it overestimates the shares of the East and West areas. However, it achieves good replication of the 3rd belt. Households choosing the 3rd belt are possibly indifferent to accessibility to employment, but care about proximity to motorways and railway stations (table I.7).

The use of the distance to the centre in model 3 (RMSE 5.42%) improves market shares replication in comparison with model 2. It underestimates to a lesser extent the share of the city centre (difference -9.3%) but it underestimates the share of the 3rd belt. These results show the weaknesses of such a simplistic definition of accessibility when we want to perform simulations. It cannot capture the variations of the perceived accessibility. The relative accessibility values are probably not representative (table I.7). The use of the time by car to the city centre (model 4), which is theoretically better than the linear distance, increases the performance of the replication (RMSE of 3.97%). Still the most problematic zone is the centre, but the underestimation is decreasing in comparison to previous models. The absence of the land-use component seems to limit the performance of this indicator.

The next two indicators integrate the land-use component of accessibility. However, the RMSE of model 5 is 6.43%, the second highest between the analysed models. It mostly fails to replicate the share of the city centre (underestimation of 11.4%) and the share of the East Areas (overestimation of 6.6%). The cumulative opportunities measure shows its weaknesses in capturing residential location preferences. This is because it gives high values of accessibility but small variation to central areas and the inverse in the periphery (table I.7). Contrary to the cumulative opportunities, the potential accessibility measure gives very good replicated market shares. Model 6 gives the best results in terms of data replication. The RMSE (1.57%) is the best between the analysed accessibility measures. These results show that the inclusion of the land-use component with this method is superior to the cumulative opportunities method (model 5). The potential indicator with a composite generalised time is performing the best and represents the best accessibility measure from a theoretical point of view.

Table I.8 *Capacity to replicate the observed market share of the models using all accessibility measures*

	Observed	Base model (no accessibility)		Transport Infr. Proximity		Dist. to centre		Time to centre		Cumm. Opp. car		Pot. Acc. Composite	
		Model 1		Model 2		Model 3		Model 4		Model 6		Model 10	
		Rep	Diff	Rep	Diff	Rep	Diff	Rep	Diff	Rep	Diff	Rep	Diff
Centre (Lyon - Villeurbanne)	55.0%	37.8%	-17.3%	39.8%	-15.3%	45.8%	-9.3%	47.8%	-7.3%	43.7%	-11.4%	53.1%	-1.9%
East Areas	11.0%	11.9%	0.9%	15.7%	4.7%	15.7%	4.7%	14.0%	3.0%	17.6%	6.6%	13.9%	2.8%
West Areas	11.1%	15.4%	4.3%	17.1%	5.9%	15.3%	4.2%	14.4%	3.3%	15.9%	4.8%	10.8%	-0.4%
2nd Belt	8.0%	12.8%	4.8%	11.9%	3.9%	11.5%	3.5%	10.2%	2.2%	10.3%	2.3%	7.3%	-0.7%
3rd Belt	14.8%	22.0%	7.2%	15.6%	0.8%	11.7%	-3.1%	13.6%	-1.2%	12.5%	-2.3%	14.9%	0.1%
		RMSE	8.87%	RMSE	7.83%	RMSE	5.42%	RMSE	3.97%	RMSE	6.43%	RMSE	1.57%

Table I.9 *Capacity to replicate the observed market share of the models including systematic taste variation and potential accessibility measures*

	Observ.	Dist. centre * Nb Indiv.		Dist. centre. * HH Status.		Time centre * Nb Indiv.		Time centre. * HH Status.		Cumm. Opp. * Nb Indiv.		Cumm. Opp. * HH Status.		Pot. Acc. Composite * Nb indiv.		Pot. Acc. Composite * HH Status	
		Model 3 - 1		Model 3 - 2		Model 4 - 1		Model 4 - 2		Model 5 – 1		Model 5 - 2		Model 6 - 1		Model 6 - 2	
		Rep.	Diff.	Rep.	Diff.	Rep.	Diff.	Rep.	Diff.	Rep.	Diff.	Rep.	Diff.	Rep.	Diff.	Rep.	Diff.
Centre (Lyon - Villeurbanne)	55%	46.3%	-8.8%	46.9%	-8.2%	48.4%	-6.6%	48.9%	-6.1%	43.8%	-11.3%	44.1%	-10.9%	54.2%	-0.9%	54.4%	-0.7%
East Areas	11%	15.5%	4.5%	15.1%	4.1%	13.7%	2.7%	13.4%	2.4%	17.8%	6.7%	17.7%	6.7%	13.3%	2.3%	13.1%	2.1%
West Areas	11%	15.1%	4.0%	14.6%	3.4%	14.1%	3.0%	13.6%	2.5%	15.8%	4.6%	15.5%	4.3%	10.6%	-0.6%	10.3%	-0.8%
2nd Belt	8%	11.0%	3.1%	11.0%	3.0%	9.8%	1.9%	9.8%	1.8%	10.1%	2.1%	9.9%	1.9%	7.4%	-0.5%	7.4%	-0.6%
3rd Belt	15%	12.0%	-2.8%	12.5%	-2.3%	14.0%	-0.9%	14.3%	-0.5%	12.6%	-2.2%	12.8%	-2.0%	14.5%	-0.3%	14.8%	0.0%
		RMSE	5.11%	RMSE	4.69%	RMSE	3.58%	RMSE	3.26%	RMSE	6.38%	RMSE	6.17%	RMSE	1.14%	RMSE	1.07%

The addition of household systematic taste variations in the model increases the performance of the replication of the observed market shares, confirming the findings of the previous chapter on their significant contribution. In all cases, the employment status of the household's head gives better results than the number of individuals for all the tested accessibility measures. The systematic taste variation by employment status interacted with the potential accessibility measure gives the best replication of the observed choices of the households (table I.9).

5. Summary and discussion

The objective of this paper is to give some answers to the question of whether the type of accessibility measure can influence the estimation results of a residential location choice model. The literature analysis has shown that there is an abundance of accessibility measures but their influence on residential choice behaviour is not always empirically identified (Blijie, 2005). For our analysis, we selected various measures, from simplistic to complex ones. Households' systematic taste variation for accessibility introduced by interacting household's characteristics with accessibility. We analysed the models' results and their ability to replicate observations.

The main result of this study is that accessibility remains a significant variable in residential location choice models when we apply global indicators, meaning the indicators that give a relative accessibility value for all locations. Their inclusion in the model increases significantly the quality and corrects for the counterintuitive sign of social housing and housing price variables. With regard to the debate around the significance of accessibility in residential location choice, we take the view that transport and land-use modellers must include some sort of accessibility measure in their models (Eliasson, 2010; Lee et al., 2010). The omission of accessibility could lead to erroneous results.

Proximity to transport infrastructure measures cannot capture the effect of the accessibility but only the local effect that can be negative, as our and other studies (De Palma et al., 2005) demonstrate. Global accessibility measures are preferable and justified because they explain utility differences between alternatives and their relative attractiveness for a residential choice. As we have shown, all global accessibility measures give consistent results, the parameters have stable signs and the selected household characteristics have a same positive sensitivity. When possible, and especially if the model is developed for simulation purposes, the use of the potential accessibility should be preferable. The application of cumulative opportunities measure should be made with care due to sensitivity to the threshold definition. The land-use component is poorly included into the cumulative accessibility measure in comparison to the potential accessibility (Geurs and van Wee, 2004). The analysis of the performance of the models to replicate the observations proved the superiority of the potential accessibility. This is because it represents better the relative importance and the variation of accessibility (table I.7) and minimises systematic errors between observed and simulated choices. Last, the inclusion of households' characteristics is not only useful from an analysis point of view but it ameliorates the replication of the data. This is because the interaction terms capture some systematic taste variation of the households' preferences.

From a policy point of view, our analysis shows the importance of the land-use and individual components of the accessibility. Decision makers should take into consideration that residential policies focusing on the supply of activities around residential areas could be more efficient at attracting new households. While it is important to have fast transport modes to reach different destinations, residents search activities near their residential location. In our case study, more than half of the moved households chose a location in central areas, showing their attractiveness. Additionally, policy makers should pay attention not to penalise sensitive groups like students or workers with no stable job, which value greatly good accessibility. Accessibility changes can affect differently the households depending on their socio-demographic characteristics, as our and other previous works highlight (Guo, 2004; Guo and Bhat, 2007).

The current work can have various extensions for future studies. We only focus on the employment opportunities, as a general proxy for various land-use opportunities. Further analysis regarding the impact of different activities other than employment is important. Empirical analyses focusing on this issue should try to resolve problems of multicollinearity of accessibilities to different activities. In addition, the accessibility measure applied is limited in terms of transport modes. The integration of more modes might be important if we want to analyse the impact of new, active and environmentally friendly transport modes such as bicycle or walk on residential choices. Last, it is interesting to analyse the temporal and spatial transferability of these results. Our future works will focus on this last issue.

Chapter II

Do new and relocating firms have different preference for accessibility?

1. Introduction

Cities are competing each other to attract new businesses and to increase their economic and general attractiveness. Nevertheless, what attracts businesses to an area is complex and depends on both location attributes and firm characteristics. One of the most important factors for a firm location decision is accessibility, highlighted by location choice theories (Alonso, 1964; Jacobs, 1969; Marshall, 1890; Von Thünen, 1842; Weber, 1909). Accessibility reflects the ease with which a location can be reached from different places, by different agents, i.e. workers, clients, suppliers, and distributors, using different transport modes. Accessibility depends both on the transport infrastructure and the spatial distribution of agents and opportunities (Geurs and van Wee, 2004). While empirical works addressing the impact of accessibility on location choices are numerous, few papers have studied the differences in the valuation of accessibility between new and relocating businesses. New and relocating businesses do not place the same importance on location attributes (Duranton and Puga, 2001; Holl, 2004a; Manjón-Antolín and Arauzo-Carod, 2011). One can assume that accessibility is one of these attributes and that different firms appreciate accessibility differently.

As considerable investments are made in transport infrastructures and land-use development aiming to increase the spatial accessibility of places, knowledge on how accessibility can affect locally the location choices of economic establishment is imperative for correct policy instruction. This paper investigates at a micro-level whether new and relocating establishments appreciate accessibility differently, using as case study the urban area of Lyon in France. The results confirm the assumption that new and relocating establishments appreciate accessibility differently. These differences are strongly related to the economic sector and the activity of the firm.

The rest of the paper is structured as follows. Section 2 focuses on the role of accessibility in location choices of new and relocating establishments. Section 3 presents the study area, the data and the applied method and the different variables and their measures. Section 4 presents the obtained results. Section 5 summarises the main findings along with the conclusions of the paper.

2. Differences between new and relocating establishments: do they appreciate accessibility differently?

When firms choose a location, either for first implantation or for relocation purposes, they evaluate a variety of location attributes based on their needs. One of the most important attributes is the accessibility of the location. A location must provide some sort of transport infrastructure in order to reduce the cost for workers, clients, suppliers and distributors (Ellison et al., 2010).

The very first theoretical works of urban economists on location choice determinants of economic activities highlight the importance of accessibility. The bid-rent theory, developed by the works of Von Thünen (1842), reveals the role of accessibility on the spatial distribution of economic activities. Since then, other economic theories have implicitly included accessibility as a location choice factor through a transport minimisation cost process (Weber, 1909), through centrality (Christaller, 1933), localisation (Marshall, 1890) or urbanisation (Jacobs, 1969). Accessibility and transport infrastructure are considered henceforth as traditional explanatory location factors with positive effect. Highly accessible areas with well-developed transport infrastructures can potentially minimise transport costs and relative risks for suppliers (input), distributors (output), labour force (production factor) and clients (profit). They also increase the potential market access helping firms to be more specialised and to exploit better economies of scale (Holl, 2012; Maroto and Zofío, 2016) creating cost efficiencies. However, the theory does not distinguish the importance of accessibility between new and migrating establishments.

Accessibility is a multicomponent concept and differences of firms' preferences can vary among its components which are (i) the transport system, (ii) the spatial distribution of land-use, (iii) the individual characteristics (Geurs and van Wee, 2004). A new firm can be more sensitive to accessibility to the local population (spatial distribution) while a relocating one can be more sensitive to transport infrastructure (transport system). These differences can vary across economic sectors (individual characteristics) of the economic establishments. While accessibility has been the focus of some studies (Bodenmann, 2011; de Bok and Van Oort, 2011), the existing literature focuses either on creations (Baptista and Mendonça, 2010; Buczkowska and de Lapparent, 2014) or relocations (Bodenmann and Axhausen, 2012; Nguyen et al., 2013; Van Dijk and Pellenbarg, 2000). Few studies compared new and relocating establishments. Manjón-Antolín and Arauzo-Carod (2011) examined the location choices of the industrial sector in Catalonia and found that either the location choices of new and relocating establishments have different sensitivity on common location attributes or that they take into account completely different determinants. Duranton and Puga (2001) observed that in France, new firms prefer diversified and central areas where the urbanisation effects are strong, while relocating enterprises prefer rural areas where the fixed costs are lower and specialisation externalities are stronger. The connection between these preferences for urbanisation/specialisation and accessibility is intuitive. Central areas with diversified economies are locations that offer high accessibility. Rural, often industrialised and specialised locations are areas with low accessibility. In line with this idea, Holl (2004a) found for Portugal that new establishments prefer local accessibility, migrating establishments prefer locations with good connections to the national market while both are attracted to transport corridors with migrations showing a greater preference for the proximity to motorways.

Both creations and relocating establishments search for profit maximising locations (Barrios et al., 2006). Nevertheless, we can distinct three major differences between them; (i) the local information they dispose, (ii) the life cycle stage and (iii) the “eco-system”. When an establishment is relocating within a geographical area, it disposes key local information about the economic environment thanks to a previous experience in the same area. Consequently it can make an informative evaluation of the different location possibilities (Manjón-Antolín and Arauzo-Carod, 2011) and make the best possible choice for its economic activity. On the contrary, a new business, which enters a new territory, misses this local information, which can be obtained only by experience. This is why a new firm will not take many risks. Regarding the life stage of a firm, a creation and a relocation of an economic establishment are two distinct events of the life course of a firm during which its needs can evolve (Holl, 2004a). A location choice made at some point back in time can now be suboptimal, pushing to search for a new one, due to factors regarding the location or the establishment itself (Van Dijk and Pellenbarg, 2000). Thus, a migrating firm is searching for a location better than the actual one (Nguyen et al., 2013). Last, when a firm is already in a geographical area it has an "eco-system" (Moore, 1993), meaning the established network with workers, suppliers and clients from a spatial perspective. This dependency on the pre-existing network poses a restriction; relocating firms do not migrate far from the previous location because it acts as a pull factor (Bodenmann and Axhausen, 2012; Van Dijk and Pellenbarg, 2000).

Using this framework to analyse potential differences in terms of preference for accessibility, we can form the following assumptions. A new firm entering a geographical area should be more risk averse but should have more freedom with no particular attachments. It will try to locate in an area where it can minimise all the potential risks with no a priori attachments. Therefore, accessibility should be in general positive but its effect should vary in terms of magnitude between creations and relocations. New firms are expected to localise themselves in central areas where accessibility is high and risks are low. On the contrary, a relocating firm can take more risks (Holl, 2004a) but has more constraints because of its "ecosystem" and its life stage. Businesses usually relocate because the current premises or the current location do not cover their needs but they choose areas which are not far from their previous location (Van Dijk and Pellenbarg, 2000), their established "eco-system".

In this paper, we provide some empirical evidence on the difference of the effect of accessibility between new and migrating economic establishments for various economic sectors. In that respect, we have developed a location choice model at a micro-level having as a case study the Lyon urban area in France. While the majority of previous works address location choices of economic at a country (Frenkel et al., 2001), region (Holl, 2004a) or commune level (Manjón-Antolín and Arauzo-Carod, 2011), this work is considering an intraurban area as a spatial unit. Knowledge at such a detailed level of analysis can highlight heterogeneities of location attributes emerging locally (Beaudry and Schiffauerova, 2009; Holl, 2004b).

3. Study area, data and model specification

3.1. The Urban Area of Lyon

The study focuses on the Lyon urban area, which is located in the southeast part of France. It is the second largest metropolitan area in France after Paris in economic and population terms. The urban area has surface area of about 3.3 thousand square km and, in 2011, had a population of 1.85 million people. It is considered as a dynamic economic area with an international character due to the proximity of the city to Italy, Switzerland and Germany (Rosales-Montano et al., 2015). The Gross Domestic Product of the metropolitan area in 2011 was almost 73 billion euros (Eurostat), which places the urban area among the 25 top European metropolitan regions in terms of total gross production. Despite the deindustrialisation process of the latest years, Lyon remains one of the most industrialised areas of France (Carpenter and Verhage, 2014). Its economy has a tertiary role, which has been reinforced in the latest years. This diversity and strength of the local economy situates Lyon between the most dynamic European metropolitan cities of this size like Cologne, Turin, Dublin, Helsinki, etc.

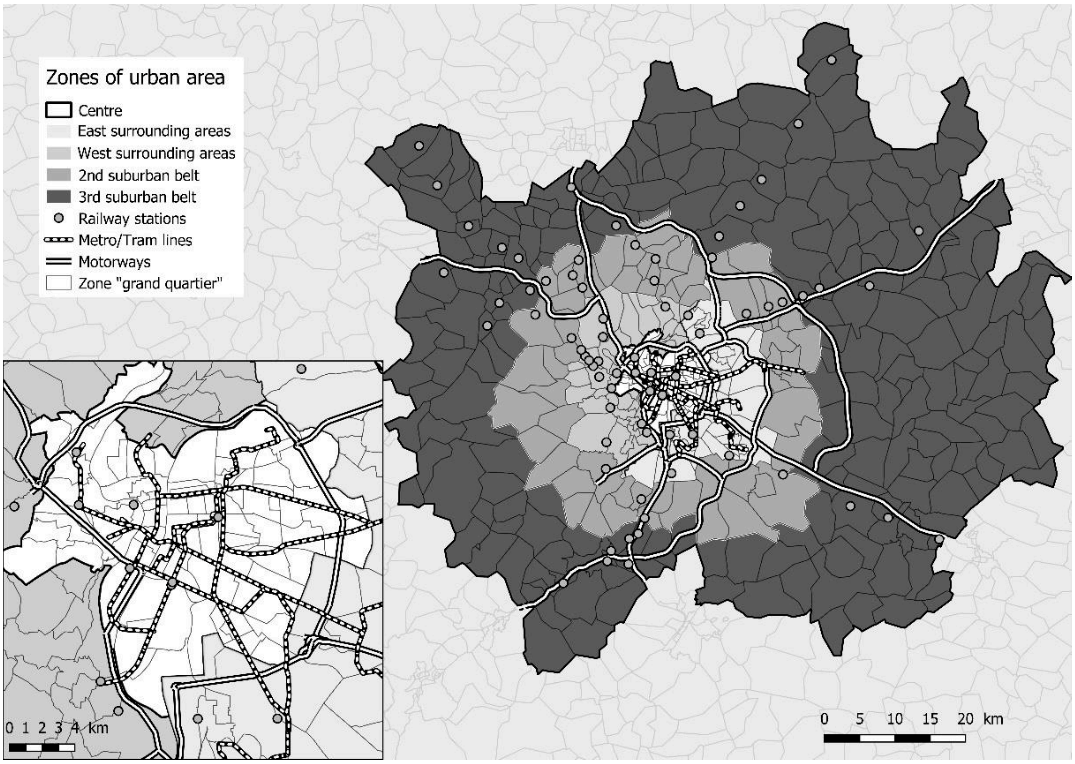


Figure II.1 *The zones and the transport infrastructure of the urban area of Lyon.*

In total, the urban area had more than 850,000 jobs (142,500 establishments - self-employed excluded) in 2011 (INSEE[3], databases of SIRENE[4] and general population census). More than 43% were concentrated in the area's central municipalities (Lyon-Villeurbanne – Centre in figure II.1) which covers less than 2% of the surface of the total analysis area. During the period 2005 – 2011, the net employment increase was almost 6% and the number of firms increased by almost 17%. Despite the economic crisis of the period, the urban area maintained its economic attractiveness thanks to its diversified economy and the growth of the tertiary sector and liberal professions. Furthermore, the local economic policy was favouring the

[3] Institut National de la Statistique et des Etudes Economiques - French National Institute of Statistics and Economic Studies
[4] The economic establishments register database

entrepreneurship with the creation of poles of competitiveness and innovation in the 90s (Rosales-Montano et al., 2015), which boosted the creation of small enterprises. All in all, findings from this paper can help understanding the behaviour of firms in a city of a medium-large size like Lyon.

3.2. Creations and migrations of economic establishments

This work is principally based on the SIRENE database enriched by other datasets from various sources. SIRENE is a disaggregated database that contains all the economic establishments in France and it is provided by the INSEE.

We use the SIRENE database for two time periods, the analysis year of 2011 and the comparison year of 2005. The use of the same database of two years allows us to identify the establishments created or migrated during this period. We define as created an establishment that was not present in the study area in 2005 but is in 2011, so it is created during the period 2005-2011. A relocating establishment is defined as an establishment which was present in the study area both in 2005 and 2011, but whose postal address has changed during this period (Nguyen et al., 2013; Pellenbarg et al., 2002). The postal address of an establishment is the address where its economic activity takes place. This provides an identification of migrations even between very close locations. The majority of firms relocate inside the municipality of origin (Bodenmann and Axhausen, 2012).

These definitions of creations and relocations are more appropriate for small firms with few establishments in our study area. Large multisite firms have different internal procedures with regards to location choices (Pellenbarg et al., 2002), which go beyond the scope of this paper. This is why we are focusing only on firms with one or two establishments in 2011. Restricting the analysis on the one or two establishments firms, we can identify migrating establishments with precision. A disadvantage of this method is the non-identification of inbound firms migrating from other regions of the country and deaths of firms during the analysis period 2005-2011. The former ones are considered as creations while the latter ones are beyond the scope of this paper. If a firm had an economic activity outside the study area before the analysis period and relocates into the study area, our method would identify its economic establishment as a creation. However, it is expected, based on our analysis framework, that migrating firms from outside the study area should behave similarly to the newly established firms since they do not have any previous experience, attachment or knowledge of the area. The advantage of this method is its transferability. It can be applied to any time period and any location for which data is available.

The SIRENE database contains information for each economic establishment like the economic activity, the postal address, the size in number of employees etc. Previous studies showed that different sectors have different location choice preferences (Bodenmann and Axhausen, 2012). In order to group firms into economic sectors we use the classification of the INSEE as a departure point. To decrease the number of sectors and to have more homogenous groups in terms of location choices we merge some groups together and recreate some others using a bottom-up approach based on the detailed activity code and the function of the firm (Table II.1). We decompose the Business services into Front Office and Back Office services. This distinction aims at reflecting the firms' needs for face-to-face interactions and the presence of structural differences linked to the degree of final demand orientation of Business Services (Ota and Fujita, 1993).

Table II.1 *Classification of economic establishments by INSEE and modifications*

Classification of INSEE	Modifications
Manufacturing	-
Production and distribution of electricity, gas, etc.	Grouped to back office services
Production and distribution of water	Grouped to back office services
Construction	-
Wholesale and retail activities	Divided to retail and wholesale
Transports and storage	Divided to back office (the majority) and front office services
Information and communication	Divided to back office and front office services (the majority)
Finance and insurance	Grouped to Finance, Insurance and Real Estate (FIRE)
Real Estate	Grouped to Finance, Insurance and Real Estate (FIRE)
Specialised, technical and scientific activities	Divided to back office and front office services (the majority)
Services and activities of support and administration	Divided to back office and front office services
Health	-

The creation and the relocation rate differs between economic sectors (table II.2). We should note that relocation rates might be underestimated because if an establishment was created and has relocated during the analysis period, this establishment is only counted as a creation. Real Estate and Construction activities have the highest rate of creations during the analysis period. These two sectors are interdependent since Construction boosts Real Estate operations and vice-versa. On the contrary, the Manufacturing and Health sectors have the weakest creation rate with only 34% creations. Bodenmann and Axhausen (2012) found for Switzerland that manufacturing activities leave cities. The mean relocation rate during the analysis period is around 11% and it differs between sectors. The most "mobile" sectors are the Front Office services and Health with 14% of relocations. This difference can be related to the difficulty of the migration (specialised premises) or to the relation of the establishment with a specific location (clients).

Table II.2 *Creation and relocation of establishments by sector*

Sector	Creations during 2005-2011	Creation rate	Migrations during 2005-2011	Migration rate
Manufacturing	2,395	34%	755	11%
Construction	6,788	52%	1,460	11%
Wholesale	3,935	47%	1,078	13%
Retail	6,039	48%	911	7%
FIRE	6,050	54%	1,249	10%
Front Office Services	9,981	48%	2,940	14%
Back Office Services	4,130	49%	989	12%
Health	4,304	34%	1,741	14%

Source: SIRENE database, authors' calculations

The majority of the created and the relocating establishments are establishments with few employees. Between 87.3% to 97.7% of the created establishments and 68.3% to 95.9% of the relocating ones, have five employees or less, with the majority of them having zero employees; only the owner is operating the establishment. As a general pattern, the created establishments are smaller than the relocating ones, which is expected. When an establishment is created, at the beginning usually only the owner is operating the business and later he recruits employees based on his needs. On the contrary, the size of migrating firms can vary significantly. This variation can be caused by the smaller sample by economic sector, especially for Manufacturing (see table II.2). The distribution is different not only between firm events but also between economic sectors. The sectors with the largest establishments (more than 20 employees) both of created and relocating are the sectors of Manufacturing (3.9% and 9,3%), Wholesale (1,4% and 6,8%) and Back Office (2,8% and 10,5%).

Table II.3 *Distribution of the establishments' size included into the analysis by sector and creation/relocation*

Sector	0 employees		1-2 employees		3-5 employees		6-9 employees		10-19 employees		20-49 employees		50+ employees	
	Creations	Migrations	Creations	Migrations	Creations	Migrations	Creations	Migrations	Creations	Migrations	Creations	Migrations	Creations	Migrations
Manufacturing	61.6%	40.8%	15.2%	13.6%	10.5%	13.9%	5.1%	10.9%	3.6%	11.5%	2.6%	6.2%	1.3%	3.0%
Construction	70.3%	49.7%	19.4%	17.9%	6.1%	12.9%	2.2%	7.5%	1.1%	6.6%	0.6%	4.2%	0.3%	1.2%
Wholesale	79.0%	47.4%	10.5%	16.5%	4.9%	11.1%	2.2%	8.7%	1.9%	9.5%	1.0%	4.7%	0.4%	2.0%
Retail	72.9%	65.8%	15.8%	18.7%	6.8%	9.2%	2.6%	3.6%	1.2%	1.4%	0.4%	1.1%	0.1%	0.2%
FIRE	83.5%	60.2%	11.1%	19.7%	3.2%	9.1%	1.2%	4.5%	0.7%	3.6%	0.3%	2.2%	0.1%	0.6%
Front Office Services	80.0%	55.3%	10.9%	16.1%	4.4%	11.4%	2.1%	6.3%	1.3%	5.7%	0.9%	3.5%	0.4%	1.7%
Back Office Services	73.8%	50.5%	12.8%	13.2%	5.6%	10.4%	2.6%	6.1%	2.4%	9.2%	1.7%	6.9%	1.0%	3.6%
Health	90.6%	84.8%	4.7%	9.3%	1.9%	1.8%	0.8%	1.2%	0.8%	1.3%	0.7%	1.1%	0.4%	0.5%

Source: SIRENE database, authors' calculations

3.3. Modelling the location choices of economic establishments

Modelling the location choice of a firm or an establishment involves various choices, ranging from the dependent variable, which can be the jobs, the firms or the establishments (Buczkowska and de Lapparent, 2014) to the modelling method. There are principally two methods, the discrete choice models, which model location choices of economic establishments and the count data models which model the attraction of locations (how many establishments are choosing a location) (Arauzo-Carod et al., 2010). Even though the two approaches mostly have the same theoretical bases, like the profit maximisation process and are relying on likelihood maximisation method for the estimation of the parameters (Alamá-Sabater et al., 2011), we use the discrete choice approach given that in our analysis we are working with establishment's data and searching how an establishment is choosing its location.

One of the most fundamental principles of the discrete choices is the McFadden's (1974) random utility maximisation principle applied to firms as a profit maximisation process. Carlton (1983) demonstrated through an empirical study in the USA that in fact the profit maximisation problem for a firm is a variant of McFadden's random utility maximisation model, applied by McFadden for the households' location choice model. A firm evaluates all the available alternative location possibilities (perfect information) and then chooses the location which maximises its profits (Arauzo-Carod et al., 2010; Holl, 2004a). In this framework, the profit Π_{in} for a firm n and a location i is composed by a deterministic observable part π_{in} and a random unobservable term ε_{in} (equation II.1) (Barrios et al., 2006):

$$\Pi_{in} = \pi_{in} + \varepsilon_{in} \tag{II.1}$$

Π_{in}: Profit of establishment n at location i
π_{in}: Deterministic part of the profit
ε_{in}: The error term

Where the deterministic part of the profit is a vector of the alternative locations' attributes:

$$\pi_{in} = \sum_{k=1}^{K} \beta_k X_{ink} \tag{II.2}$$

π_{in}: Deterministic part of the profit
K: the number of variables
β_k: parameter to be estimated
X_{ink}: value of variable for establishment n at location i

Making the assumption that the error terms ε_{in} are independently and identically distributed (IID) with type 1 extreme value distribution (McFadden, 1974), the probability of choosing a location takes the logit form:

$$Pn(i) = \frac{e^{\pi_{in}}}{\sum_{j \in C_n} e^{\pi_{jn}}} \quad (II.3)$$

π_{in}: *Deterministic part of the profit at i*
π_{jn}: *Deterministic part of the profit of all alternatives in* C_n
C_n: *the choice set*

The Multinomial logit model is strongly criticized especially for modelling spatial choices. One of its key properties is known as the IIA (Independence of Irrelevant Alternatives) property. This property means that the probability to choose between two options is independent of the rest non-chosen alternatives (Ben-Akiva and Lerman, 1985). However, this assumption is highly unlikely to be valid in a spatial context where alternatives can be correlated (Ibeas et al., 2013), as space is artificially divided. However, it remains an attractive method due to the ease of computation and the traceability of the results.

In our case, the alternatives are the zones of the study area and we estimation a separate model for each of the eight analysed economic sectors. Because we search for potential differences between new and relocating establishments, we assume that we have two market segments by economic sector, the creations and the relocations. Thus, equation II.2 becomes:

$$\pi_{in} = \sum_{k=1}^{K} \alpha_{k1} X_{ink} \delta_1 + \sum_{k=1}^{K} \alpha_{k2} X_{ink} \delta_2 \quad (II.4)$$

π_{in}: *Deterministic part of the profit*
K: the number of variables
α_{k1}, α_{k2}: *parameters to be estimated*
X_{ink}: *value of variable for establishment n at location i*
δ_1, δ_2: *Binary indicators of created or relocated establishments*

Where δ_1 and δ_2 are indicators which take the values 0 or 1 depending on whether the observation is a creation or a relocation. Using this joint estimation we are able to compare the results (Ben-Akiva et al., 2015) between creations and relocations. In order to verify that this division is explaining better the location choices and to justify our choices, we estimate a model where we do not make the distinction between the firm events and we compare the results using the likelihood ratio test (Ben-Akiva and Lerman, 1985).

The study area is divided into 431 zones[5] and we use the full choice set for the estimation of the parameters. Even though our focus is accessibility, we integrate other location attributes highlighted by the location choice theory of firms in order to control for their effect. The selected locational attributes are classified into four groups: accessibility and market trade-off, location externalities, social environment and institutional factors. Table II.4 presents the summary statistics of all the independent variables. Such historical data are not available for each year of the analysis period (2005-2011). To assure that the explanatory variables are

[5] We are following the zoning system of INSEE based on the « *grand quartier* » zones, which is essentially a census breakdown

exogenous to location choices, we estimate all the variables for the initial year, using data from 2005.

3.4 The independent variables

3.4.1. Accessibility and market trade-off.

This group of variables is the centre of our analysis. We want to capture the maximum information regarding the preference for accessibility using various indicators. At the same time, we want to measure the traditional trade-off between accessibility of a location and price. Especially for the relocating establishments, this trade-off includes the distance to the previous location.

In order to measure accessibility, we use three types of measures: (i) the proximity to transport infrastructure, which captures the effect of the presence of an infrastructure, (ii) the preference for centrality, and (iii) the potential accessibility indicator, which combines the ease to travel and the spatial distribution of the population. The first and second type of accessibility are easily observable by the firm like proximity or not to a transport infrastructure or central area or not (de Bok and Van Oort, 2011). The third measure is less intuitive but is a more comprehensive indicator of accessibility.

The proximity to transport infrastructures includes the stations of PT (metro, tramway and railway) and the motorway. The proximity to these infrastructures is measured as a binomial variable, which takes the value 1 when this type of infrastructure is present into the zone. This measurement method avoids potential correlation with other accessibility indicators.

In order to capture the preference for central areas we introduce a set of dummy variables. We divide the area into five greater areas so that we have: (i) the central zone composed by the municipalities of Lyon and Villeurbanne, (ii) the eastern surrounding zone which is considered as areas with low skilled workers, (iii) the western surrounding zone which is considered as rich areas with highly skilled workers, (iv) the 2nd suburban belt, and (v) the 3rd suburban belt. In that way, we capture not only the preference for the central areas, but also the preferences, if any, between those different greater areas. Previous studies have highlighted the importance of centrality of a location (Dubé et al., 2016; Elgar et al., 2009).

The potential accessibility (Geurs and van Wee, 2004), measures the population potentially attracted to a given location and can be interpreted as a proxy for the potential market. It allows to verify the role of the interaction level of the economic activities with the population. While there can be some sectors which appreciate accessibility to a pool of workers than to population, we observed that accessibility to workers and to population are highly correlated. Thus, from an estimation point of view, the results would be similar either using accessibility to workers or to population. The form of the accessibility to population using a negative exponential impedance function, is given by the equation II.5:

$$A_j = \sum_i P_i \, e^{-\beta T_{ij}} \qquad \text{(II.5)}$$

A_j: Accessibility to population of location j
P_i : Population in location i
β: parameter to be estimated
T_{ij}: generalised time from i to j

The parameter β of the equation II.5 is inferred from the calibration of the distribution step of a four-stage model using local household travel survey. This parameter reflects the sensitivity of individuals to make a trip given the distance between the origin and the destination. We tested several combinations of specific population segments based on socio-demographic profiles, but the results did not vary significantly between the definitions. For the sake of simplicity and comparability, we chose to use the total population for all economic sectors.

In the study area, most of the trips are made by car and by Public Transport (PT). A calculation of an accessibility using composite generalised times allows to take into account more than one transport mode in an area. For the aggregation of the generalised times, we applied a method developed by Bhat et al. (1999). The result is a composite generalised time, which combines generalised times by car and PT (equation II.6). When PT is not available, the generalised time by PT is equal to +∞. This formulation is chosen because it is theoretically consistent with the concept of accessibility as the benefit of using the transport system; the opportunities represent the utility and the time the disutility to reach those opportunities. The increase of transport solutions must be associated with a decrease of the disutility to reach opportunities (Handy and Niemeier, 1997).

$$T_{ij} = \left(\frac{Tcar_{ij}}{1 + \frac{Tcar_{ij}}{Tpt_{ij}}}\right) \qquad \text{(II.6)}$$

T_{ij}: Composite generalised time from i to j
$Tcar_{ij}$: Generalised time for car
Tpt_{ij}:Generalised time for PT

As the traditional bid-rent theory states, there is a trade-off between the accessibility and the price of a location. In order to capture this, we introduce the price per square meter for different types of premises. It is expected to have a negative influence *ceteris paribus* and higher importance for relocating establishments.

For migrating establishments, the remoteness from the previous location is included as an independent variable. It is measured as the Euclidian distance between the previous location, i.e. zone, and the zone of relocation. This distance is expected to have negative influence on the location choice.

The proximity to transport infrastructure indicators concern the year 2005 and they are calculated using GIS based on spatial data provided by the IGN (National Geographic Institute of France). The generalised peak times by car and PT were estimated by a four-step transport model developed at the LAET (Laboratoire, Aménagement, Economie, Transport - Transport, Urban Planning, Economics Laboratory). The data of the household travel survey of 2006 is used for the calibration of the model and the estimation of the generalised times, because it is the closest to the initial year of our analysis. The population data comes from the general census (RGP) of 2005 provided by the INSEE. The data for the premises' prices comes from the Callon database (Callon, 2005), which gives an average price per square metre for offices, boutiques, warehouses, and industrial premises for 2005. This data is combined with data of apartment sales from the database of the notaries of France (Perval) of the same year.

3.4.2. Agglomeration externalities.

Agglomeration externalities or external economies are important determinants for a location choice of a firm, highlighted by the neoclassical approach (Arauzo-Carod et al., 2010). They arise when firms increase their productivity because of the proximity to other firms without any direct financial exchanges. These agglomeration externalities can be divided in two different types, the localisation and the urbanisation effects (Glaeser et al., 1992). Known as MAR[6] externalities, the localisation externalities emerge from the concentration of an economic sector to a specific geographical area. It is considered as a positive location externality because proximity between firms can favour the labour market pooling, input/output sharing and knowledge spill over (Ellison et al., 2010). It increases the performance of firms and reduces the risk for the implementation of new ones. In empirical applications, localisation effects are measured either by using the location quotient by economic sector or by the density of employment or firms (Beaudry and Schiffauerova, 2009).

[6] Marshall-Arrow-Romer externalities as formalised by (Glaeser et al., 1992)

In this work, after testing for different formulations, we use the density of firms by location and by sector which gives the best and the most consistent results. The density LOC is given by equation II.7:

$$LOC_{js} = \frac{nb_{js}}{E_j} \qquad \text{(II.7)}$$

nb: the total number of establishments
s: economic sector
E: surface in km²
j: zone

Localisation effects can be a proxy for accessibility components but are not captured by our accessibility indicators. These effects are closely related to accessibility (de Bok and Van Oort, 2011; Melo et al., 2016) but from a broader firm-to-firm influence point of view without considering the influence of the infrastructure (de Bok, 2007), because it concerns only the physical proximity for companies in the same industry. Additionally, localisation externalities can be a proxy for good suppliers accessibility due to "shared input effects" (Marshall, 1890).

The diversity externalities (Jacobs, 1969) are the result of the colocation of diverse economic sectors into a geographic area (Beaudry and Schiffauerova, 2009). There are economic sectors, which value more the diversity than the density of a location while others search for specialised locations. The diversity effects can be measured by the Gini coefficient or the Hirschman-Herfindahl index (HHI) (Beaudry and Schiffauerova, 2009). We have opted for the HHI, modifying it as HHI'=1-HHI (see equation II.8) in order to have more intuitive results (values between 0 and 1, 1 the most diverse).

$$HHI'_j = 1 - \frac{\sum_s {D_{sj}}^2}{(\sum_s D_{sj})^2} \qquad \text{(II.8)}$$

D: number of jobs
s: economic sector
j: zone

As shown by Duranton and Puga (2001), localisation effects are expected to be stronger for relocating establishments and diversity effects stronger for creations. For the calculation of those two indicators, we rely on the data of the 2005 SIRENE database.

3.4.3 Social environment.

Even though studies do not include social environment variables very often, we argue that it can influence the location choice of a firm. Firms which offer high quality services and need face-to-face contact are expected to choose high revenue areas (Elgar et al., 2009). These areas are more attractive from a human capital and from a market potential point of view especially for certain economic sectors. On the contrary, firms might avoid areas with high concentration of low-income households. Such an area can have negative local effects, due to possible social problems that can affect the productivity of a firm, while small disposable revenue to spend can impact consuming oriented activities. For some economic sectors, such deprived areas can be attractive because they provide low-skilled workers.

Table II.4 *Summary statistics of the variables used in the model*

Group	Variable	Description	Mean	Median	Std. dev.	Min	Max
Accessibility and market trade-off	**Potential accessibility to population**	Accessibility to general population	197	116	170	8	595
	Motorway	Presence of a motorway into the zone	0.29	0.00	0.45	0.00	1.00
	Metro/Tramway station	Presence of a metro or a tramway station into the zone	0.13	0.00	0.34	0.00	1.00
	Railway station	Presence of a railway station into the zone	0.22	0.00	0.42	0.00	1.00
	Centrality	Preference for the central zone (Lyon Villeurbanne) in comparison to the other 4 zones	0.18	0.00	0.39	0.00	1.00
	Premise's price for industrial use	The mean price per square meter for premises designated for industrial use (€ 2005)	572	510	266	227	1,468
	Premise's price for commercial use	The mean price per square meter for premises designated for commercial use (€ 2005)	1,471	1,338	629	469	4,391
	Premise's price for office use	The mean price per square meter for premises designated for office use (€ 2005)	969	895	306	429	2,075
	Premise's price for storage use	The mean price per square meter for premises designated for storage use (€ 2005)	455	406	231	291	1,180
Agglomeration externalities	**Manufacturing localisation**		0.16	0.02	0.40	0.00	3.19
	Construction localisation		0.17	0.03	0.30	0.00	2.00
	Wholesale localisation	Density of establishments of each sector group (number of establishments by 100km^2)	0.21	0.02	0.48	0.00	5.47
	Retail localisation		0.48	0.02	1.39	0.00	12.91
	FIRE localisation		0.31	0.02	0.90	0.00	7.73
	Front Office Services localisation		0.47	0.03	1.27	0.00	8.73
	Back Office Services localisation		0.12	0.02	0.22	0.00	1.26
	Health localisation		0.35	0.02	0.80	0.00	6.68
	Hirschman-Herfindahl index (Diversity)	The inversed HHI	0.81	0.84	0.084	0.36	0.91
Social environment	**% Q1**	The percentage of the population in the 1st quantile of revenue	0.15	0.13	0.08	0.05	0.58
	% Q5	The percentage of the population in the 5th quantile of revenue	0.21	0.21	0.10	0.02	0.55
Institutional factors	**Economic Activity zone**	Presence of an Economic Activity Zone	0.10	0.00	0.30	0.00	1.00

In our study, we are taking into account the effect of the social environment by introducing into the model the percentages of the population belonging to the 1st quantile (the poorest) and the 5th quantile (the richest) of the revenue of the whole study area. We have used the DGI database (INSEE), which gives the distribution of fiscal revenues of each zone, for the year 2005. We expect that some sectors will be positive to the 5th quantile while most of the sectors will be negatively influenced by the 1st quantile.

3.4.4. Institutional factors.

One of the latest advances of the location theory is the understanding that firms are making choices in an environment which is not static, because of agents who are external to the firm like the government. Public incentives can have great influence to the location choices of firms (Barrios et al., 2006). To account for the role of the macro-agent (public authorities, government) we have integrated a binary variable for the Economic Activity Zones (Zones d'Activité Economique – ZAE). A ZAE is a designated geographic area of concentration of economic activity, organised and constructed by a public or private developer, who rents or sells the land and the premises to companies willing to locate their businesses in these areas (Cerema, 2014). The identification of the Economic Activity Zones was made through personal research, combining different sources, since there is no official registry. For the centre, the east and west surrounding areas, we found the data from the official site of the metropolitan territorial authority (data.grandlyon.com). For the rest, we used information from reports of the INSEE, from websites for businesses (e.g. lyon-entreprises.com, zonedactivite.com) and from the communes which dispose a ZAE, in order to find its exact location of the zone. It is expected that this variable has a positive effect. In general, these areas are located near transport axes and create localisation or diversity effects so controlling for its effect is crucial.

4. Modelling results

We estimate a different model for each economic sector combining creations and relocations as described in equation II.4. This set of models is called Model I. Then, we include the distance to the last location for the relocating firms and we re-estimate the models. This set of models is called Model II and for the analysis of the results, we only focus on the variables concerning the relocations, since the variables of the created establishments are not changing.

In this paper, we only present the results of elasticities calculated for the variables. For the quantitative variables we calculate the mean point elasticities and for the categorical variables we estimate the mean pseudo-elasticities (equations II.9 and II.10) (Washington et al., 2011). The result of the mean point elasticity suggests the mean effect that an increase of 1% of this variable will have. Accordingly, the result of the mean pseudo-elasticity is the mean effect of a categorical variable when it changes from 0 to 1. Thus, the results between the quantitative and categorical variables cannot be directly compared. The detailed results of the models can be found in the appendix.

Mean point elasticity	$E_k = \frac{\sum_{i=1}^{I}(1 - P_{in})\,\hat{\beta}_k\, X_{ink}}{I}$	*E_k: Mean elasticity for quantitative variable k* *I: The number of establishments* *P_{in}: The probability of establishment i choosing the location n* *$\hat{\beta}_k$: The estimated parameter for k* *X_{ink}: The value of k for i at n*	(II.9)
Mean pseudo-elasticity	$E'_k = \frac{\sum_{i=1}^{I}\frac{P_{in}(i\mid X_{ink}=1) - P_{in}(i\mid X_{ink}=0)}{P_{in}(i\mid X_{ink}=0)}}{I}$	*E'_k: Mean pseudo-elasticity for categorical variable k* *P_{in}: The probability of establishment i choosing the location n* *X_{ink}: The value of k for i at n which can take the values 0 and 1*	(II.10)

4.1. New and relocating economic establishments choose their locations differently

First, we want to test our assumption that created and relocated establishments have different location choice behaviours. For this reason we use the likelihood ratio test (Ben-Akiva et al., 2015). The restrained model is the model with all the establishments without distinction between new and relocating ones, and the unrestricted model, is the model with market segments (creations and relocations). The test is given by equation II.11. If the test is rejected, then market the segmentation model is better, meaning that the new and relocating establishments do choose their locations differently.

$LRT = -2\,(L(\hat{\beta}) - L_{\delta}(\hat{\beta}))$	*$L(\hat{\beta})$: Log-likelihood of the restricted model* *$L_{\delta}(\hat{\beta})$: Log-likelihood of the unrestricted model* *The test is x2 asymptotically distributed with K degrees of freedom (K: the difference of variables between the two models)*	(II.11)

Table II.5 presents the log-likelihood of the models and the results of the likelihood ratio test. For all economic sectors, the segmentation between new and relocating establishments is justified.

Table II.5 *Likelihood ratio test for segmentation between created and relocated establishments*

Sector	Manufacturing	Construction	Wholesale	Retail	FIRE	Front Office	Back Office	Health
Log of likelihood (no segmentation)	-18,332	-48,146	-28,461	-39,722	-40,166	-69,098	-29,691	-34,450
Log of likelihood (market segments)	-18,290	-48,059	-28,387	-39,637	-40,144	-69,014	-29,650	-34,423
Likelihood ratio test	84	173	147	171	44	169	81	53
Significant (K=14)	At 99%	At 99%	At 99%	At 99%	At 99%	At 99%	At 99%	At 99%

The results show that new and relocating establishments choose their location choices differently. However, this result can be related to the size difference between the creations and the relocations observed previously (table II.3). In order to control this, we estimate additional models excluding the very small establishments (less than two employees). These additional models confirm our intuition that the differences do not come from the difference in size, since the Likelihood ratio test is still significant.

4.2. Different events, different preferences

Adjusted rho-squared varies between 0.040 and 0.142 (appendix, tables II.A1-II.A3). These values may seem small but it is due to the high number of alternatives (Guevara and Ben-Akiva, 2013). The high variation of the values means that some sectors have more heterogeneous location choices than others. The addition of the distance to the last location for the relocating firms increases the quality of the model for all sectors. This fact reveals the importance of the last location for the relocating firms. In the rest of the chapter we analyse the results of the estimated mean elasticities and mean pseudo-elasticities in detail by variable groups and by firm event highlighting the differences between creations and relocations. The results are shown in tables II.6 to II.9.

4.2.1. Accessibility and market trade-off.

The influence of accessibility to population is significant for creations of all sectors, while for relocations we observe an important variation of the significance and the elasticity depending on the sector. For Manufacturing, the parameter is positive for creations and negative for migrations, meaning that only creations consider accessibility to population as a positive location attribute. For Construction Wholesale and Back Office, accessibility to population is significant and positive only for the creations. For migrations while it is not significant and positive in Model I, it becomes negative in Model II. This means that for relocations of these sectors, accessibility has marginal impact, while the distance to the last location has a very strong effect. FIRE and Front Office services appreciate good accessibility to population, with high elasticities. Elasticity for migrations in Model I is higher than for creations. However, in Model II, accessibility to population becomes non-significant. An explanation might be the relation between migrating distance and accessibility; the smaller the migrating distance the smaller the relative difference between the accessibilities of the two locations (the previous and the current). Thus, what is important is the distance to the previous location. Other studies have showed that these types of economic activities appreciate good accessibility to population (Bodenmann and Axhausen, 2012; de Bok, 2007; de Bok and Van Oort, 2011). Those two sectors are the only ones of which the migrations are not only sensitive to accessibility, but they have higher levels of elasticity when we do not take into account the distance to the previous location. Last, accessibility to population for Retail and Health is important for both creations and relocations, with a higher elasticity for the creations. In model II, we can make the same observation as for the previous sectors. Accessibility becomes non-significant for Retail and negative for Health. For these sectors, accessibility is valuable for both events, but more important for creations.

Proximity to transport infrastructure has an overall significant positive effect. Proximity to motorways is the most important transport infrastructure for almost all sectors. This observation is consistent for both events. The pseudo-elasticity varies between 29% and 111%. An exception is the FIRE sector, which has slightly higher preference for railway stations and Health, which appreciates more proximity to metro/tramway stations. When we compare creations and relocations, a pattern emerges. Migrations have higher sensitivity to the proximity to motorways than the creations. Only migrations of Health are indifferent to motorways. This observation is in line with previous studies that found the same sensitivity at a national level (Holl, 2004a). It seems that the same holds at the intraurban level. Last, proximity to metro/tramway and railway stations does not provide such a clear pattern. It seems that creations of Manufacturing, Construction, Wholesale, Back Office and Health sectors value more proximity to such PT infrastructure than migrations, while it is the opposite for Retail, FIRE and Front Office.

In general, the premise's price has a negative influence but it is not always the case. Confirming our assumptions, the negative elasticities are higher for the migrations, meaning that migrations are more sensitive to the land value. The integration of the distance to the last location (model II) reinforces the observed negative impact. In some cases, for Manufacturing, Construction, Wholesale, Retail, Back Office and Health the price has a positive or non-significant parameter. Other studies found the same contradictory results (Bodenmann and Axhausen, 2012). With respect to previous observations concerning accessibility, it seems that when the establishments are not searching for locations with high accessibility, price parameter cannot be balanced leading to inconsistent parameters (Bodenmann and Axhausen, 2012).

Regarding the preference for centrality or for other greater areas, there is a difference between creations and migrations. Relocations of Manufacturing, Construction Wholesale and Back Office sectors have a smaller preference for central areas than the creations and they prefer better the western areas and the 2nd belt. This result is in line with the results of Duranton and Puga (2001). For FIRE, Front Office establishments we observe that central areas are more important for migrations. For Retail and Health, we do not observe any difference between creations and relocations.

These results regarding accessibility variables confirm our initial hypotheses. Creations appreciate more areas with good PT infrastructure and high accessibility to population while migrating establishments appreciate more proximity to motorways and avoid high priced areas. The negative and not intuitive effect of accessibility to population in Model II can be caused by omitted variables or by the strong influence of the distance to last location. Other studies have found similar results. Elgar et al. (2009) found for Toronto that relocating office firms value positively accessibility. However, its parameter became negative when they added the distance to the last location. De Bok (2007) included in the location choice model the distance to the previous location and found a non-significant effect of potential accessibility for relocating establishments. He argues that this effect is potentially caused by the accessibility measure. Another study from De Bok and Oort (2011) found in the South Holland (a Dutch region) that the logsum accessibility for trips to work for relocating firms does not have a significant effect for the new location choices, after accounting for the distance to the previous location. It seems that for relocating firms, the distance to the previous location dominates the decision for the new location (de Bok and Van Oort, 2011; Elgar et al., 2009; Sweet, 2014), due to possible risk aversion of the firm (Van Dijk and Pellenbarg, 2000). A location away for the last location where the firm has already developed its "*eco-system*" involves some sort of risks especially for the mobility habits of clients, labour and suppliers.

4.2.2. Agglomeration externalities.

In accordance with the theory, results show that localisation always have a positive significant effect (Barrios et al., 2006) for both creations and relocations. Migrating establishments from most of the sectors have a higher preference for localisation than the creations. Nevertheless, migrating establishments from Construction and Manufacturing show a marginally smaller preference for localisation.

The results of the economic diversity of the location do not provide us a clear picture. For some sectors, diversity is more important for creations (Manufacturing, Front Office, Back Office and Health), in line with (Duranton and Puga, 2001). For the other sectors, diversity is more important for migrations. Thus, we cannot conclude on the direction of the difference between creations and relocations.

Table II.6 *Elasticities for creations and migrations in Manufacturing and Construction*

	Manufacturing				Construction			
	Model I		Model II		Model I		Model II	
Variable	Creations	Migrations	Creations	Migrations	Creations	Migrations	Creations	Migrations
Localisation [quantitative]	0.22*	0.18*	0.22*	0.19*	0.24*	0.07*	0.24*	0.09*
Diversity [quantitative]	1.29*	1.05*	1.29*	1.07*	1.02*	2.09*	1.02*	1.96*
Accessibility pop. [quantitative]	0.28*	-0.01	0.28*	-0.86*	0.22*	0.17	0.22*	-0.55*
Motorway [categorical]	0.52*	1.07*	0.52*	0.88*	0.40*	0.42*	0.40*	0.34*
Metro/Tramway [categorical]	0.14*	-0.06	0.14*	0.03	-0.10*	-0.22*	-0.10*	-0.14
Railway Station [categorical]	0.06	0.13	0.07	0.10	0.11*	0.23*	0.11*	0.20*
Centre (Reference)								
Eastern Areas [categorical]	0.11	0.71*	0.11	0.58*	0.15*	0.78*	0.15*	0.74*
Western Areas [categorical]	0.09	0.06	0.09	-0.05	0.09	0.15	0.09	0.06
2nd Belt [categorical]	0.33*	0.36*	0.33*	0.59*	0.14	0.64*	0.14	0.96*
3rd Belt [categorical]	-0.27*	-0.58*	-0.26*	0.33	-0.37*	-0.33*	-0.37*	1.01*
Q1 % [quantitative]	-0.14*	-0.29*	-0.13*	-0.34*	0.07	-0.21*	0.07	-0.31*
Q5 % [quantitative]	-0.32*	-0.20	-0.30*	-0.09	-0.31*	-0.45*	-0.31*	-0.33*
ZAE [categorical]	1.00*	1.15*	1.01*	1.04*	0.83*	1.14*	0.83*	1.07*
Premise's price [quantitative]	-0.06	-0.34	-0.06	-0.58*	0.00	0.11	0.00	-0.20
Distance last loc [quantitative]				-1.97*				-2.01*

** significant parameter at 95%*

Table II.7 *Elasticities for creations and migrations in Wholesale and Retail*

	Wholesale				Retail			
	Model I		Model II		Model I		Model II	
Variable	Creations	Migrations	Creations	Migrations	Creations	Migrations	Creations	Migrations
Localisation [quantitative]	0.14*	0.14*	0.14*	0.16*	0.47*	0.02	0.48*	0.36*
Diversity [quantitative]	1.32*	1.56*	1.32*	1.42*	0.17	0.04	1.19*	1.84*
Accessibility pop. [quantitative]	0.34*	0.21	0.34*	-0.48*	1.27*	0.39*	1.01*	0.05
Motorway [categorical]	0.45*	1.11*	0.45*	1.00*	-0.08*	0.48*	0.05	0.33*
Metro/Tramway [categorical]	0.18*	0.21*	0.18*	0.27*	0.03	-0.03	0.03	0.03
Railway Station [categorical]	0.25*	0.01	0.25*	-0.05	0.25*	0.32*	0.51*	0.12
Centre (Reference)								
Eastern Areas [categorical]	0.05	0.50*	0.05	0.69*	0.21*	0.24*	0.32*	0.48*
Western Areas [categorical]	0.05	0.22	0.05	0.23	0.25*	0.25*	0.07	0.41*
2nd Belt [categorical]	0.03	0.38*	0.02	1.28*	0.14	0.12	-0.09	0.60*
3rd Belt [categorical]	-0.58*	-0.60*	-0.58*	0.85*	-0.27*	-0.30*	-0.57*	2.13*
Q1 % [quantitative]	-0.09	-0.35*	-0.09	-0.42*	-0.05	-0.01	-0.17*	-0.30*
Q5 % [quantitative]	0.07	-0.03	0.07	-0.05	0.06	0.01	0.10*	0.12
ZAE [categorical]	1.11*	2.16*	1.11*	2.10*	0.47*	0.49*	0.49*	0.92*
Premise's price [quantitative]	0.08	0.00	0.08	-0.17	-1.12*	0.62	-1.31*	-1.77*
Distance last loc [quantitative]				-1.77*				-1.80*

** significant parameter at 95%*

Table II.8 *Elasticities for creations and migrations in R.E Finance & Insurance and Front Office*

	R.E Finance & Insurance				Front Office			
	Model I		Model II		Model I		Model II	
Variable	Creations	Migrations	Creations	Migrations	Creations	Migrations	Creations	Migrations
Localisation [quantitative]	0.25*	0.44*	0.25*	0.47*	0.46*	0.81*	0.46*	0.83*
Diversity [quantitative]	0.78*	0.25	0.79*	0.24	1.21*	1.46*	1.21*	1.35*
Accessibility pop. [quantitative]	0.62*	0.77*	0.62*	-0.09	0.73*	0.84*	0.73*	0.00
Motorway [categorical]	0.35*	0.39*	0.35*	0.35*	0.29*	0.54*	0.29*	0.44*
Metro/Tramway [categorical]	0.31*	0.34*	0.31*	0.37*	0.30*	0.41*	0.30*	0.39*
Railway Station [categorical]	0.38*	0.52*	0.37*	0.46*	0.27*	0.38*	0.27*	0.27*
Centre (Reference)								
Eastern Areas [categorical]	-0.13*	-0.05	-0.13*	0.25	-0.32*	-0.21*	-0.32*	0.38*
Western Areas [categorical]	0.19*	0.16	0.19*	0.29*	0.06	0.11	0.07	0.37*
2nd Belt [categorical]	-0.25*	-0.36*	-0.25*	0.21	-0.44*	-0.58*	-0.44*	0.25
3rd Belt [categorical]	-0.61*	-0.76*	-0.61*	0.09	-0.74*	-0.87*	-0.73*	0.49*
Q1 % [quantitative]	-0.21*	-0.44*	-0.21*	-0.51*	-0.18*	-0.34*	-0.19*	-0.46*
Q5 % [quantitative]	0.67*	1.07*	0.67*	1.08*	0.76*	0.91*	0.74*	0.88*
ZAE [categorical]	0.93*	0.91*	0.93*	1.07*	0.89*	1.06*	0.89*	1.36*
Premise's price [quantitative]	-0.32*	-1.36*	-0.31*	-1.80*	-0.95*	-1.78*	-0.95*	-2.35*
Distance last loc [quantitative]				-1.60*				-1.45*

** significant parameter at 95%*

Table II.9 *Elasticities for creations and migrations in Back Office and Health*

	Back Office				Health			
	Model I		Model II		Model I		Model II	
Variable	Creations	Migrations	Creations	Migrations	Creations	Migrations	Creations	Migrations
Localisation [quantitative]	0.09*	0.17*	0.09*	0.18*	0.26*	0.26*	0.26*	0.22*
Diversity [quantitative]	0.76*	0.26	0.76*	0.42	0.40*	0.32	0.39*	0.38
Accessibility pop. [quantitative]	0.27*	0.05	0.27*	-0.87*	0.34*	0.31*	0.34*	-0.61*
Motorway [categorical]	0.43*	0.54*	0.43*	0.43*	0.14*	0.01	0.13*	-0.05
Metro/Tramway [categorical]	-0.04	-0.09	-0.04	-0.04	0.25*	0.07	0.25*	0.07
Railway Station [categorical]	0.17*	-0.11	0.17*	-0.09	0.10*	0.14*	0.10*	0.02
Centre (Reference)								
Eastern Areas [categorical]	0.01	0.44*	0.02	0.31*	-0.16*	-0.08	-0.16*	0.06
Western Areas [categorical]	-0.06	0.03	-0.06	-0.09	0.32*	0.32*	0.32*	0.39*
2nd Belt [categorical]	-0.03	0.24	-0.03	0.55*	-0.07	-0.16	-0.07	0.44*
3rd Belt [categorical]	-0.49*	-0.63*	-0.49*	0.30	-0.57*	-0.71*	-0.57*	0.30
Q1 % [quantitative]	0.11	-0.39*	0.11*	-0.36*	0.06	0.42*	0.05	0.44*
Q5 % [quantitative]	-0.34*	-0.47*	-0.34*	-0.34*	-0.11	0.29*	-0.12	0.35*
ZAE [categorical]	0.98*	1.35*	0.98*	1.17*	0.45*	0.41*	0.45*	0.49*
Premise's price [quantitative]	0.17*	0.03	0.17*	-0.05	0.25	0.07	0.22	-0.04
Distance last loc [quantitative]				-1.93*				-2.03*

** significant parameter at 95%*

4.2.3. Social environment.

The differences between new and relocating establishments, regarding the appreciation of the social environment, are more accentuated than the location externalities. Sensibility is not only different in terms of elasticity but also in terms of the effect (positive/negative). For FIRE and Front Office, relocations avoid deprived neighbourhoods and they prefer rich areas more than the creations. Migrations of Manufacturing, Construction, Wholesale and Back Office sectors avoid both deprived and rich areas more than creations. For Health, both rich and deprived areas seem to affect positively the location choice of relocations. We can conclude that the social environment is a location attribute that is appreciated differently from creations and relocations. It is an attribute which needs time to be evaluated and an establishment, which has a previous knowledge of the local area can make a better choice for its economic activity.

4.2.4. Institutional factors.

For almost all sectors, relocating establishments have higher preference for Economic Activity Zones than the creations. This means that relocating establishments are in position to take advantage of these zones and all the positive effects they offer.

5. Conclusions

The object of this article is to highlight the differences of the location choices between newly created and migrating economic establishments, focusing on accessibility variables. In theory, those two types of location choices should be different in terms of preferences for accessibility because they are at a different firm life stage. They do not have the same local experience based information and they do not have the same "*eco-system*" constraints.

Our results show that the location choices of new and relocating establishments differ. The hypothesis that creations are more sensitive to accessibility to population and relocations more sensitive to transport infrastructure generally holds. Table II.10 summarises the results. We can classify the sectors in three groups based on the differences of preferences for accessibility. In the first group, we have the sectors of Manufacturing, Construction, Wholesale and Back Office. These activities have a routine role and they are production oriented. Creations are searching for locations with good accessibility in order to minimise all the potential risks. However, the migrating establishments of these sectors are searching for areas with better proximity to transport infrastructure. These locations offer good access to national and international markets (Holl, 2004a). In the second group we have the FIRE and Front Office. These activities are high-order services and require daily face-to-face interaction and information exchange (Shearmur and Alvergne, 2002) and accessibility is important for their activity. This is why migrating establishments, which have the previous experience of the area, choose areas that have better accessibility in general (better accessibility to population/better proximity to transport infrastructure). In the third group, we have the Retail and Health. The economic activity of each of these sectors is specific and depends highly on client interaction.

However, we can observe that the creations of these sectors have higher preference for accessibility to population in line with our initial hypotheses. For the other indicators there is no clear pattern.

Last, regarding only the migrations, they are highly sensitive to the distance of the previous location, independently of the economic sector confirming the findings of the literature (Elgar et al., 2015; Holl, 2004a; Manjón-Antolín and Arauzo-Carod, 2011). Firms want to be in proximity to their established "*eco-system*". Most of the times, the inclusion of the distance to the last location corrects for the non-intuitive sign of the premises' price. As long as the new location is near to the previous one, accessibility has marginal or even negative effect because the two locations offer relatively close accessibility levels. This explains why accessibility has a negative influence for relocating firms, a result found in other studies as well (de Bok, 2007).

Table II.10 *Synthesis of the observed differences in preferences for accessibility for all types of accessibility and all sectors*

Accessibility type		Manufacturing	Construction	Back Office	Wholesale	Front Office	FIRE	Retail	Health
		Sector							
Accessibility to population	Between Creations and Migrations, it influences more the location choice of:	C+	C+	C+	C+	M+	M+	C+	C+
Motorway		M+	M+	M+	M+	M+	M+	M+	C+
Metro/Tramway		C+	M-	ND	M+	M+	M+	ND	C+
Railway Station		ND	M+	C+	C+	M+	M+	M+	M+
Eastern Areas		M+	M+	M+	M+	C-	C-	M+	M-
Western Areas		ND	ND	ND	ND	ND	C+	ND	ND
2nd Belt		M+	M+	ND	M+	M-	M-	ND	ND
3rd Belt		M-	C-	M-	M-	M-	M-	M-	M-

Note: C: Creations, M: Migrations, ND: No difference, +: Positive influence, -: Negative influence. Lecture example: For the Manufacturing sector, the influence of accessibility is more important for Creations (C) in a positive way. For the Construction sector, proximity to Motorway influences more the location choices of the Migrating (M) establishments in a positive way (+). For the Back Office, areas at the 2nd belt do not have any different effect (ND) between Creations and Migrations.

The results of this work are important from a policy perspective. Cities are investing massively in transport projects in order to attract businesses. Evidence from this paper shows that local authorities must pay attention to those policies. For some sectors, public transport infrastructure is more important for creations while for others it is more important for relocations. Motorways are more important for relocations. For most non tertiary sectors, accessibility to population is more important for creations. Transport policies aiming to attract new firms could provoke relocations from nearby locations decreasing the economic activity inside the same urban area. Tailor-made policies with clear objectives are needed to attract the economic activities that the policies are aiming.

While in this paper we focused on the determinants attracting economic establishments (pull factors), the location attributes that drive economic activity away (push factors) is equally important. In future research, we want to turn our interest on the location determinants that contribute to firms' deaths and migrations in order to have a clearer picture on those pull-push factors.

6. Appendix

Table II.A1 *Model results of manufacturing, construction and wholesale sectors*

		Manufacturing				Construction				Wholesale			
		Model I		**Model II**		**Model I**		**Model II**		**Model I**		**Model II**	
	Variable	**Coeff.**	**T stat.**	**Coeff.**	**T stat.**	**Coeff.**	**T stat.**	**Coeff.**	**T stat.**	**Coeff.**	**T stat.**	**Coeff.**	**T stat.**
Creations	**Localisation**	0.60	10.58	0.60	10.59	0.71	13.67	0.71	13.68	0.28	9.67	0.28	9.69
	Diversity	1.57	5.20	1.57	5.19	1.24	6.88	1.24	6.87	1.61	6.77	1.61	6.75
	Accessibility pop.	0.99	2.45	0.99	2.46	0.78	3.22	0.78	3.24	1.12	3.72	1.12	3.71
	Motorway	0.42	8.85	0.42	8.85	0.33	12.17	0.33	12.16	0.37	10.14	0.37	10.15
	Metro/Tramway	0.13	1.93	0.13	1.93	-0.11	-2.79	-0.11	-2.79	0.17	3.42	0.17	3.42
	Railway Station	0.06	1.32	0.06	1.33	0.10	3.57	0.10	3.57	0.22	6.27	0.23	6.29
	Centre (Reference)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Eastern Areas	0.10	1.01	0.10	1.03	0.14	2.53	0.14	2.53	0.05	0.71	0.05	0.69
	Western Areas	0.08	0.74	0.09	0.75	0.08	1.33	0.08	1.33	0.05	0.59	0.05	0.57
	2nd Belt	0.28	2.05	0.29	2.06	0.13	1.59	0.13	1.59	0.03	0.25	0.02	0.22
	3rd Belt	-0.31	-1.93	-0.31	-1.91	-0.46	-4.97	-0.46	-4.97	-0.87	-7.14	-0.88	-7.18
	Q1 %	-0.85	-1.96	-0.81	-1.87	0.39	1.67	0.39	1.67	-0.57	-1.68	-0.58	-1.71
	Q5 %	-1.51	-3.45	-1.46	-3.33	-1.63	-6.45	-1.63	-6.45	0.32	1.05	0.32	1.05
	ZAE	0.69	12.40	0.70	12.43	0.61	17.91	0.61	17.91	0.74	17.91	0.75	17.92
	Premise's price	-0.08	-0.44	-0.09	-0.50	0.00	-0.01	0.00	0.04	0.14	0.98	0.14	0.97
Migrations	**Localisation**	0.69	6.02	0.72	6.21	0.32	2.26	0.39	2.72	0.37	6.26	0.41	6.94
	Diversity	1.28	2.36	1.31	2.43	2.54	5.94	2.39	5.65	1.90	3.92	1.73	3.63
	Accessibility pop.	-0.03	-0.04	-3.32	-4.45	0.68	1.27	-2.23	-3.98	0.73	1.23	-1.72	-2.80
	Motorway	0.73	8.82	0.63	7.54	0.35	5.96	0.29	4.81	0.75	10.48	0.69	9.62
	Metro/Tramway	-0.06	-0.50	0.03	0.24	-0.25	-2.75	-0.15	-1.69	0.19	2.00	0.24	2.43
	Railway Station	0.13	1.53	0.10	1.16	0.20	3.39	0.18	2.89	0.01	0.09	-0.05	-0.70
	Centre (Reference)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Eastern Areas	0.53	3.03	0.46	2.53	0.57	4.50	0.56	4.26	0.40	2.76	0.52	3.56
	Western Areas	0.05	0.26	-0.05	-0.24	0.14	0.94	0.06	0.43	0.20	1.22	0.20	1.24
	2nd Belt	0.31	1.31	0.46	1.93	0.50	2.85	0.67	3.79	0.32	1.64	0.82	4.16
	3rd Belt	-0.87	-3.11	0.29	0.99	-0.40	-1.97	0.70	3.28	-0.91	-3.79	0.61	2.50
	Q1 %	-1.80	-2.43	-2.10	-2.77	-1.30	-2.30	-1.91	-3.32	-2.41	-3.50	-2.88	-4.15
	Q5 %	-0.95	-1.24	-0.44	-0.59	-2.26	-4.11	-1.63	-2.98	-0.12	-0.19	-0.23	-0.38
	ZAE	0.77	8.36	0.71	7.61	0.76	11.29	0.73	10.63	1.15	15.77	1.13	15.33
	Premise's price	-0.52	-1.59	-0.91	-2.79	0.18	0.76	-0.32	-1.39	0.01	0.03	-0.29	-1.06
	Distance last loc.	NA	NA	-0.27	-31.68	NA	NA	-0.28	-45.71	NA	NA	-0.24	-34.07
	Observations (segment - new)	**2395**		**2395**		**6788**		**6788**		**3935**		**3935**	
	Observations (segment - relocations)	**755**		**755**		**1460**		**1460**		**1078**		**1078**	
	Observations (total)	**3150**		**3150**		**8248**		**8248**		**5013**		**5013**	
	Likelihood zero	**-19108**		**-19108**		**-50033**		**-50033**		**-30409**		**-30409**	
	Log of Likelihood	**-18290**		**-17481**		**-48059**		**-46213**		**-28387**		**-27580**	
	Adjusted ρ^2	**0.044**		**0.087**		**0.040**		**0.077**		**0.067**		**0.096**	

Table II.A2 *Model results of retail, R.E Finance & Insurance and Front Office sectors*

		Retail				**R.E Finance & Insurance**				**Front Office**			
		Model I		**Model II**		**Model I**		**Model II**		**Model I**		**Model II**	
	Variable	**Coeff.**	**T stat.**	**Coeff.**	**T stat.**	**Coeff.**	**T stat.**	**Coeff.**	**T stat.**	**Coeff.**	**T stat.**	**Coeff.**	**T stat.**
Creations	**Localisation**	0.26	31.99	0.26	28.82	0.20	14.58	0.20	14.60	0.22	24.42	0.22	24.51
	Diversity	0.20	1.18	1.45	7.69	0.96	5.28	0.96	5.31	1.47	10.46	1.47	10.46
	Accessibility pop.	3.72	15.22	2.97	12.23	1.81	7.52	1.81	7.51	1.94	10.51	1.96	10.62
	Motorway	-0.08	-2.45	0.05	1.63	0.30	9.39	0.30	9.39	0.25	9.68	0.25	9.67
	Metro/Tramway	0.03	0.73	0.03	0.71	0.27	7.12	0.27	7.13	0.26	9.05	0.26	9.01
	Railway Station	0.22	7.01	0.41	13.43	0.32	10.86	0.32	10.85	0.24	10.47	0.24	10.35
	Centre (Reference)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Eastern Areas	0.19	3.15	0.28	4.88	-0.14	-2.11	-0.14	-2.13	-0.39	-7.45	-0.39	-7.43
	Western Areas	0.22	3.24	0.07	1.02	0.17	2.65	0.17	2.64	0.06	1.15	0.06	1.24
	2nd Belt	0.13	1.42	-0.09	-1.03	-0.29	-3.31	-0.29	-3.32	-0.59	-8.46	-0.58	-8.34
	3rd Belt	-0.31	-3.00	-0.85	-8.46	-0.93	-9.01	-0.93	-9.01	-1.33	-16.38	-1.33	-16.28
	Q1 %	-0.27	-1.09	-0.98	-3.99	-1.43	-4.32	-1.43	-4.30	-1.19	-4.61	-1.24	-4.82
	Q5 %	0.27	1.03	0.46	1.72	2.54	9.69	2.54	9.68	2.88	14.49	2.83	14.22
	ZAE	0.38	9.37	0.40	10.47	0.66	18.20	0.66	18.19	0.64	21.88	0.64	21.93
	Premise's price	-0.56	-11.76	-0.65	-12.05	-0.25	-2.36	-0.25	-2.30	-0.74	-8.81	-0.74	-8.77
Migrations	**Localisation**	0.02	0.28	0.26	8.69	0.30	10.25	0.32	10.84	0.32	20.04	0.32	20.28
	Diversity	0.05	0.12	2.25	4.45	0.31	0.79	0.30	0.77	1.78	6.70	1.65	6.21
	Accessibility pop.	1.23	2.97	0.16	0.24	2.13	4.04	-0.25	-0.46	2.07	6.09	0.00	-0.01
	Motorway	0.39	5.10	0.28	3.66	0.33	4.62	0.30	4.17	0.43	8.69	0.36	7.22
	Metro/Tramway	-0.03	-0.35	0.03	0.29	0.29	3.55	0.32	3.86	0.34	6.52	0.33	6.28
	Railway Station	0.28	4.03	0.11	1.40	0.42	6.49	0.38	5.74	0.32	7.63	0.24	5.54
	Centre (Reference)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Eastern Areas	0.22	1.47	0.39	2.72	-0.05	-0.37	0.23	1.50	-0.24	-2.54	0.33	3.37
	Western Areas	0.22	1.40	0.34	2.15	0.14	0.98	0.26	1.77	0.11	1.10	0.32	3.39
	2nd Belt	0.12	0.54	0.47	2.14	-0.45	-2.31	0.19	0.99	-0.86	-6.59	0.22	1.64
	3rd Belt	-0.36	-1.44	1.14	4.43	-1.42	-6.04	0.08	0.33	-2.01	-12.55	0.40	2.23
	Q1 %	-0.04	-0.05	-1.76	-2.81	-3.20	-3.97	-3.66	-4.50	-2.24	-4.59	-3.04	-6.21
	Q5 %	0.05	0.12	0.56	0.76	3.84	6.76	3.87	6.74	3.39	9.32	3.28	8.90
	ZAE	0.40	4.17	0.65	7.04	0.65	8.07	0.73	8.93	0.73	13.77	0.86	16.08
	Premise's price	0.33	1.37	-0.94	-5.64	-1.04	-4.57	0.00	-6.05	-1.33	-8.70	-1.76	-11.37
	Distance last loc.	NA	NA	-0.27	-31.98	NA	NA	-0.29	-35.71	NA	NA	-0.32	-49.71
	Observations (segment - new)	**6039**		**6039**		**6050**		**6050**		**9981**		**9981**	
	Observations (segment - relocations)	**911**		**911**		**1249**		**1249**		**2940**		**2940**	
	Observations (total)	**6950**		**6950**		**7299**		**7299**		**12921**		**12921**	
	Likelihood zero	**-42159**		**-42159**		**-44277**		**-44277**		**-78380**		**-78380**	
	Log of Likelihood	**-39637**		**-38658**		**-40144**		**-39184**		**-69014**		**-67247**	
	Adjusted ρ^2	**0.061**		**0.084**		**0.093**		**0.116**		**0.120**		**0.142**	

Table II.A3 *Model results of back office and health sectors*

	Variable	Back Office				Health			
		Model I		Model II		Model I		Model II	
		Coeff.	T stat.	Coeff.	T stat.	Coeff.	T stat.	Coeff.	T stat.
Creations	Localisation	0.42	3.76	0.42	3.77	0.25	13.15	0.25	13.22
	Diversity	0.93	4.23	0.93	4.23	0.49	2.37	0.48	2.33
	Accessibility pop.	0.94	3.07	0.94	3.08	1.02	3.58	1.02	3.58
	Motorway	0.36	9.95	0.36	9.96	0.13	3.41	0.13	3.39
	Metro/Tramway	-0.04	-0.92	-0.04	-0.92	0.23	4.78	0.23	4.75
	Railway Station	0.16	4.37	0.16	4.38	0.09	2.61	0.10	2.65
	Centre (Reference)	NA	NA	NA	NA	NA	NA	NA	NA
	Eastern Areas	0.01	0.21	0.01	0.22	-0.18	-2.47	-0.18	-2.44
	Western Areas	-0.06	-0.75	-0.06	-0.75	0.28	3.71	0.28	3.70
	2nd Belt	-0.03	-0.33	-0.03	-0.32	-0.07	-0.69	-0.07	-0.69
	3rd Belt	-0.68	-5.83	-0.68	-5.82	-0.84	-7.16	-0.84	-7.19
	Q1 %	0.60	1.98	0.60	1.99	0.35	1.14	0.28	0.89
	Q5 %	-1.75	-5.71	-1.75	-5.71	-0.50	-1.59	-0.52	-1.65
	ZAE	0.68	16.41	0.68	16.42	0.37	7.86	0.37	7.87
	Premise's price	0.33	2.34	0.33	2.34	0.21	1.73	0.18	1.49
Migrations	Localisation	0.96	4.00	0.97	4.01	0.24	8.06	0.20	6.58
	Diversity	0.32	0.73	0.52	1.20	0.40	1.24	0.47	1.45
	Accessibility pop.	0.17	0.27	-3.18	-4.87	0.92	2.12	-1.79	-3.93
	Motorway	0.43	5.78	0.36	4.77	0.01	0.10	-0.05	-0.86
	Metro/Tramway	-0.10	-0.96	-0.04	-0.35	0.06	0.88	0.06	0.89
	Railway Station	-0.12	-1.56	-0.10	-1.23	0.13	2.22	0.02	0.39
	Centre (Reference)	NA	NA	NA	NA	NA	NA	NA	NA
	Eastern Areas	0.37	2.47	0.27	1.77	-0.08	-0.74	0.06	0.48
	Western Areas	0.03	0.17	-0.09	-0.53	0.28	2.47	0.33	2.84
	2nd Belt	0.21	1.06	0.44	2.15	-0.18	-1.14	0.37	2.19
	3rd Belt	-0.99	-4.09	0.27	1.06	-1.22	-6.69	0.26	1.16
	Q1 %	-2.37	-3.62	-2.18	-3.31	2.39	5.40	2.51	5.48
	Q5 %	-2.29	-3.62	-1.67	-2.66	1.28	2.77	1.53	3.29
	ZAE	0.85	10.63	0.77	9.45	0.34	4.53	0.40	5.08
	Premise's price	0.05	0.18	-0.10	-0.35	0.06	0.30	-0.04	-0.19
	Distance last loc.	NA	NA	-0.27	-35.70	NA	NA	-0.44	-50.95
	Observations (segment - new)	**4130**		**4130**		**4304**		**4304**	
	Observations (segment - relocations)	**989**		**989**		**1741**		**1741**	
	Observations (total)	**5119**		**5119**		**6045**		**6045**	
	Likelihood zero	**-31052**		**-31052**		**-36670**		**-36670**	
	Log of Likelihood	**-29650**		**-28630**		**-34423**		**-31746**	
	Adjusted ρ^2	**0.046**		**0.079**		**0.062**		**0.135**	

Chapter III
The impact of accessibility on the location choices of business services. Evidence from Lyon urban area

1. Introduction

The importance of accessibility for the explanation of the location choices of economic establishments has been highlighted at a theoretical level in the very first works of urban economists on location choice determinants of economic activities. The bid-rent theory, developed by Von Thünen (1842) and extended by Alonso (1964), Mills (1967) and Muth (1969), reveals the role of accessibility on the spatial distribution of economic establishments. Other theories refer implicitly to accessibility as a location choice factor through a transport minimisation cost process (Weber, 1909), through centrality (Christaller, 1933), agglomeration (Marshall, 1890) or urbanisation (Jacobs, 1969). Accessibility and proximity to transport infrastructures are considered henceforth as traditional explanatory location attributes with positive effect.

In this paper, we search the extent to which different business services have different appreciation of the accessibility for their location choice process based on their function (Front Office, Back Office). We use an urban setting of a medium to large size European city. Evidence from a city of that size can enrich the current literature. From a transport policy perspective, quantifying the impact of accessibility can facilitate the policy decision making, design and evaluation. Different transport policies can attract or discourage certain functions of business services.

The main contribution of the article is the comprehension of the business services location choice behaviour in relation to accessibility. While accessibility has been the focus of some studies (Bodenmann, 2011; de Bok and Van Oort, 2011), existing literature focuses on business services in general (de Bok and Van Oort, 2011) or specifically on the knowledge intensive business services (Dubé et al., 2016). However, there is a basis that firms in business services sector choose their location based on their function (Duranton and Puga, 2005; Ota and Fujita, 1993), thus analysing this distinction is essential. Last, the analysis is carried out at the neighbourhood level. Knowledge in a such detailed level of analysis can highlight the heterogeneities of location attributes emerging locally (Beaudry and Schiffauerova, 2009; Holl, 2004b).

Accessibility is defined by (i) the transport system, (ii) the spatial distribution of land-use and (iii) the individual dimension (Geurs and van Wee, 2004). These components form the concept of accessibility and from the firm point of view, they influence a location choice decision jointly:

(i) Transport network had always been important for the location choices of firms. Even at the beginning of the industrial revolution, industries were looking to be located near railway stations or rivers. Today, proximity to transport infrastructures like motorways or Public Transport (PT) facilities is something that business owners take into account in their location choice decision (Mejia-Dorantes et al., 2012). This is because proximity to such infrastructures can increase potential clients and can facilitate the access of workers and other associate firms. Studies from USA and Europe confirm its importance. Transport infrastructure like motorways attracts employment from the areas around them in the USA (Duranton and Turner, 2012) and in Paris (Padeiro, 2013) while in Spain new infrastructures attract firms around them at the expense of other areas (Holl, 2004b).

(ii) The relative spatial distribution of firms, clients and workers influences the potential interaction between them. For firms we can distinguish 4 different components relevant to the spatial dimension of accessibility; the industry, the suppliers, the labour and the client levels. The industry level concerns the spatial distribution of firms in the same industrial sector or in a different one. The supplier level, even though it can be somehow related to the industry level, concerns the actual suppliers of the firm, and real interaction (not potential) is needed to be identified, which is very difficult to be measured (specific survey is needed). If suppliers are far away, this can cause increases in costs and time that can affect efficiency and profit directly. Next, the labour level concerns the spatial distribution of the active working population potentially available for the firm. Easy access to a pool of workers can increase the possibility of recruiting better-qualified stuff and can decrease commuting costs and potential risks like absenteeism. Finally, the client level concerns the spatial distribution of the potential clients, which can be the population or other firms. These clients should be able to visit the firm, if the firm offers a service in its premises, or the firm should be able to offer its services by distance. Therefore, relative proximity between clients and firms is essential but its importance can vary depending on the activity of the firm.

(iii) The individual dimension concerns firm specific activities, preferences and abilities. For firms this dimension can have two perspectives; the internal perspective from the point of view of the firm and the external perspective, which concerns all other agents external to the firm. The internal perspective influences the ability of the firm to attract labour, clients or suppliers. This ability depends on the characteristics of the firm like size, age and economic sector. Especially for the location choice process, a firm characteristic to be considered is the firm event, new creation or relocation. However, these internal characteristics should be matched with the characteristics of the agents external to the firm, the external perspective of the individual dimension of accessibility (Martín-Barroso et al., 2017). These external characteristics apply not only to workers, but also to clients and suppliers and can influence the potential relation with the firm (Martín-Barroso et al., 2017). Last, an aspect that should be considered in the individual dimension is the competition between firms and these different agents (Geurs and van Wee, 2004). Firms whose activities are in the same economic sector would potentially compete for a work force with the same qualifications as well as for the same group of clients. However, to account properly for competition effects, the study area should minimize incoming flows of workers and clients from external zones (Bunel and Tovar, 2014).

Highly accessible areas with well-developed transport infrastructures can potentially minimise transport costs for suppliers (input), distribution (output), labour (production factor) and clients (profit). In that way, it can create cost efficiencies and can be considered as a positive attribute of a location (de Bok and Van Oort, 2011). In that sense, areas that offer high accessibility are ideal for business services. For knowledge intensive business services, Dubé et al. (2016) found that proximity to central areas is very important for a non-metropolitan area in Canada. Baptista and Mendonça (2010) found similar results for Portugal. De Bok and Van Oort (2011) observed that migrating establishments of business services do appreciate accessibility and proximity to transport infrastructure in a Dutch region. In France, Buczkowska and de Lapparent (2014) focused on the Paris metropolitan area and examined the location choice of newly created firms for various sectors. They found that accessibility affect positively location choices of the Special, Scientific and Technical activities. These firms are sensitive to public transport and to the distance to motorways. Thus, we can assume that accessibility should be a key location choice factor, especially for business service activities. These activities have a high degree of final demand orientation and a high need for proximity to similar Business Services like Research and Development or Business Administration, the so called Front Office services (Ota and Fujita, 1993).

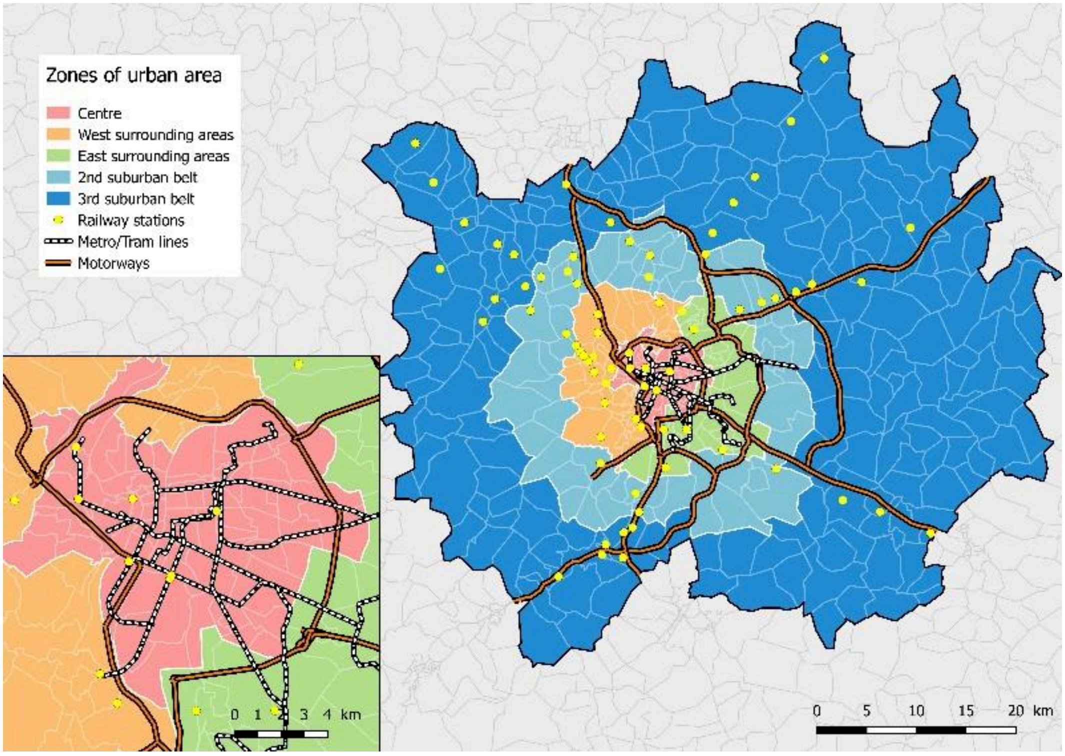
Nevertheless, during the last years, the influence of accessibility seems to be shifting. Urban areas have faced important transformations because of the dispersion of economic activities (Mejia-Dorantes et al., 2012). Some types of firms avoid high priced central areas and search for locations at periphery where rents are lower so they can increase their margin of profit. The phenomenon of the dispersion of economic activities is not independent to the changes on the transport sector. Transport was the accelerator of the rapid suburbanisation of cities during the post war era (Baum-Snow, 2007; Glaeser and Kahn M. E., 2004). These changes decreased the cost of transport for goods and people, a traditional location choice factor (Boiteux-Orain and Huriot, 2002), and gave more flexibility to firms when they are choosing a location. Taking advantage of this cost minimisation, business services are decentralising completely or only specific functions of their activities which have a more routine character in order to decrease their expenses (for land and salaries), the so called "back officing" of routine functions (Ota and Fujita, 1993). Back-office activities are services that can be provided from distance like equipment rental or call centers. In that sense, these Back Office functions should be less sensible to physical accessibility.

The rest of the paper is structured as follows. Section 2 presents the study area and section 3 the data used in our analysis. Section 4 presents in detail the applied method and presents the different variables and their measures. Sections 5 makes a summary of the data concerning the economic establishments and section 6 presents the results and the analysis of the models comparing the two subsections. Section 7 summarises the findings along with the conclusions of the paper.

2. Study area

The study area is the Lyon urban area, which is situated at the southeast central part of France. Lyon is the second largest metropolitan area in France after Paris in economic and population terms. The urban area has surface of about 3,3 thousand squared km and had a population of 1,8 million people in 2011. It is considered as a dynamic economic area with an international character due to the proximity of the city to Italy, Switzerland and Germany (Rosales-Montano et al., 2015). The Gross Domestic Product of the metropolitan area in 2011 was almost 73 billion euros (Eurostat), which places the urban area among the 25 top European metropolitan regions in terms of total gross production. Despite the deindustrialisation process of the latest years, Lyon stays one of the most industrialised areas of France (Carpenter and Verhage, 2014). Nevertheless, its economy has a tertiary role, which is reinforced during the latest years. This diversity and strength of the local economy places Lyon between the most dynamic European metropolitan cities of this size like Cologne, Turin, Dublin, Helsinki etc.

Figure III.1 *Study area and transport infrastructure*



The urban area had more than 850.000 jobs (142.500 jobs from self-employed establishments are excluded) in 2011, of which more than 43% were concentrated in the area's central municipalities (Lyon-Villeurbanne) and almost 77% inside the so-called "Greater Lyon", which is made up of the city of Lyon and some suburbs. The local economic policy is favouring the entrepreneurship with the creation of poles of competitiveness and innovation during the 90s (Rosales-Montano et al., 2015). During the period 2005 – 2011, the number of jobs has increased by almost 6% and the number of firms by almost 17%. Evidence from this article can help understanding the behaviour of firm in such urban contexts which can differ from the global European metropolitan areas like Paris (Buczkowska and de Lapparent, 2014; Padeiro, 2013) or other American cities (Sweet, 2014) on which the research is mostly focused.

3. Data sources

This work is principally based on the register of economic establishments (SIRENE database) which is a disaggregated database that contains all the companies in France. It is provided by the INSEE (Institut National de la Statistique et des Etudes Economiques - French National Institute of Statistics and Economics Studies).

We used the SIRENE database for two time periods, the analysis year of 2011 and the comparison year of 2005. This period allows us to have enough observations for creations and relocations for the model estimation. Additionally, during the same period, accessibility has improved thanks to the creation of two new tram lines in 2006 and 2009. The use of the same database in two time periods allows us to identify the firms created or relocated during this period. For identification reasons, we focus only on firms with one or two establishments in 2011. A disadvantage of this method is the non-identification of the inbound firms, which are considered as creations. However, it is expected that newcomers behave similarly to the new firms since they do not have any previous attachment in the study area. The advantage of this method is its transferability. It can be applied to any time period and any location for which there is available data.

The database contains several characteristics for each economic establishment like the economic activity, the location of the firm, the size in number of employees etc. In order to group the firms on business services based on their function, the detailed classification (NAF code) of the INSEE was used. We decompose the Business services into Front Office and Back Office services. This distinction aims at reflecting the firms' need of face-to-face interactions and the presence of structural differences, linked to the degree of final demand orientation of Business Services. In the annexe, we are presenting the categorisation of the INSEE and the retained grouping of this article.

For the estimation of the models we have used other data sources as well. For the calculation of the accessibility indicators, we used generalised travel times by car and public transport (PT) combined with the population of the area. The peak-hour generalised travel times by private vehicle and PT were estimated by a four-step transport model developed in LAET (Laboratoire, Aménagement, Economie, Transport - Transport, Urban Planning, Economics Laboratory) for the urban area of Lyon. For the calibration of the model, the data of the household travel survey of 2015 was used. Even though there might be some changes between 2011 and 2015, especially for PT with the construction of some new tram stations, they are considered marginal in terms of travel times and thus this data is applicable to our case. The population data comes from the national census for the year 2012. The other accessibility indicators, like the proximity to transport infrastructure, were calculated using Geographic Information Systems. The data for the real-estate prices comes from the Callon database, which gives an average price per square metre for offices, boutiques, warehouses and industrial premises. Thus, we were able to estimate different prices for different economic sectors. The calculation of the sectoral agglomeration and urbanisation effects was based on the SIRENE database as well. The identification of the Economic Activity Zones was made through personal research since there is no official register. Last, in order to characterise the social environment we have used the FILOSOFI database of INSEE for the year 2012, which gives the distribution of the available revenues for households of each zone.

4. Modelling the location choices

4.1. Model specification

One of the most fundamental principles of the discrete choices is the McFadden's (1974) random utility maximisation principle applied to firms as a profit maximisation process. Carlton (1983) demonstrated through an empirical study that in fact the profit maximisation problem for a firm is a variant of the McFadden's random utility maximisation model for the households. In this framework a firm is evaluating all the available location possibilities (perfect information) and then chooses the location which maximises its profits (Barrios et al., 2006). Even though some assumptions seem unrealistic (e.g. perfect information), the framework is appealing from a theoretical and computational perspective. Thus, the profit is a function of a deterministic and a stochastic part (equation III.1):

$$\Pi_{in} = \pi_{in} + \varepsilon_{in} \quad \text{(III.1)}$$

Π_{in}: Profit of establishment i at location n
π_{in}: Deterministic part of the profit
ε_{in}: The error term

Making the assumption that the error term ε_{ij} is independently and identically distributed (IID) with type 1 extreme value distribution (McFadden, 1974), the probability of choosing a location takes the logit form:

$$P_{in} = \frac{e^{\pi_{in}}}{\sum_{j \in C_n} e^{\pi_{jn}}} \quad \begin{array}{l} P_{in}\text{: The probability of individual } i \\ C_n\text{: The choice set of alternative zones} \end{array} \qquad \text{(III.2)}$$

where the deterministic part π_{in} is given by equation III.3:

$$\pi_{in} = \sum_{k=1}^{K} \beta_k X_{ink} \quad \begin{array}{l} \pi_{in}\text{: Deterministic part of the profit} \\ K\text{: The number of variables} \\ \beta_k\text{: Parameter to be estimated} \\ X_{ink}\text{: Value of variable for individual } i \text{ at location } n \end{array} \qquad \text{(III.3)}$$

Among the assumptions of the logit model, the violation of the hypothesis IID (Identical and Independent Distribution) of residuals that generates the IIA property (Independence of Irrelevant Alternatives) is the most important problem especially in a spatial context. However, as the literature states, the Multinomial logit model stays an attractive method due to the ease of computation and the traceability of the results which stay consistent (De Palma et al., 2005). In our case, the area zoning system is divided in 431 zones[7] which has been chosen is order to reduce zone similarities. We are also using the full choice set for the estimation of the parameters. The selected zone by the establishment takes the value 1 and all the others the value 0. The developed model focuses on the accessibility variables. However, there is a need to integrate other location attributes mentioned in the location theory in order to control for their effect and to have consistent results. These attributes can be classified in four groups: accessibility and market trade-off, location externalities, social environment and institutional factors. Based on the correlation matrix, there is no serious multicolinearity between the variables. The highest one is between the accessibility to population and the land value, which is 0.6.

4.2. Variables and measures

4.2.1. Accessibility and market trade-off

In order to measure the accessibility, we have selected two types of measures: 1) the proximity to transport infrastructure, which captures the effect of the presence of an infrastructure, and 2) the potential accessibility indicator, which combines the ease to travel and the spatial distribution of the population. The first type of accessibility is easily observable by the firm like proximity or not to a transport infrastructure (de Bok and Van Oort, 2011). The second measure is less intuitive but is a more comprehensive indicator of accessibility.

[7]The zones of *« grand quartier »* is used. A *« grand quartier »* is a grouping of census zones inside the same commune. Its size in terms of population varies strongly but in general, a commune need to have more than 10.000 population to be divided. For example, a commune of 20.000 population would be divided in 2 or 3 *« grand quartiers »*.

The potential accessibility indicates the population that can reach potentially the firm's location. We have tested several combinations of specific population segments based on socio-demographic and socio-economic profiles, but the results did not vary significantly between them. Thus, for the sake of simplicity and comparability of the results, we have chosen to use the general population. A general form of the measure for origin and destination locations i and j respectively the accessibility for population P and travel time by a mode of transport t is given by Hansen (1959):

$$A_j = \sum_i P_i \, e^{-\beta Tc_{ij}} \qquad \text{(III.4)}$$

A_j: *Accessibility to population of zone j*
P_i: *Population at i*
β: *Parameter to be calibrated*
t_{ij}: *Composite generalised time from i to j*

The parameter β is estimated using local data of the trip behaviour of individuals and reveals the effect of the time on the probability to make a trip. When we have multiple transport modes serving an area, one should consider aggregating between modes in order to calculate a combined accessibility. For the aggregation of the generalised times, we applied a method developed by Bhat et al. (1999). The result is a composite generalised time, which combines generalised times by car and PT (equation III.5). The generalised time by car is the reference time for all pairs of OD because it is always available. When PT is not available, the generalised time by PT is equal to +∞. The idea is when both car and PT serve an Origin-Destination (OD) pair, the generalised time should be less than the fastest mode. With more mobility options, it is easier to commute between ODs. Thus, accessibility should be higher. This formulation is chosen because it is theoretically consistent with the concept of accessibility as the benefit of using the transport system; the opportunities represent the utility and the time the disutility to reach those opportunities. The increase of transport solutions must be associated with a decrease of the disutility to reach opportunities (Handy and Niemeier, 1997). There exist other methods to calculate composite generalised times, but the chosen method gave the best results in the residential location choice context of Lyon. A limit of the aggregation of car and PT times is that we lose the relative influence of each mode. Possibly accessibility by car and by PT have different effect on location choices. However, the inclusion of both accessibilities in the same model creates collinearity issues.

$$Tc_{ij} = \left(\frac{Tcar_{ij}}{1 + \frac{Tcar_{ij}}{Tpt_{ij}}}\right) \qquad \text{(III.5)}$$

Tc_{ij}: *Composite generalised time from i to j*
$Tcar_{ij}$: *Generalised time for car*
Tpt_{ij}: *Generalised time for PT*

The result of the potential accessibility to population can be interpreted as a proxy for the market potential, which allows verifying the role of the proximity of the economic activities to population. This means that firms who need face-to-face contact should be more sensitive to accessibility. However, in general, accessibility should be considered as a positive location attribute in all cases as we described in section 2.

To represent the proximity to transport infrastructures, we include binary variables that take the value of 1 if there is a PT (metro, tramway or railway) station or a motorway section present in the zone, else 0. We have not used a continuous measure, like the distance to the motorway, because we want to capture only the local effect of the infrastructure.

Previous studies have highlighted the importance of centrality of the location (Dubé et al., 2016; Elgar et al., 2009). In our case, in order to capture this preference for central areas we have introduced a set of dummy variables. We have divided the area in 5 greater areas (see figure III.1) where we have: (i) the central zone composed by the municipalities of Lyon and Villeurbanne, (ii) the eastern surrounding zone which is considered as areas with low skilled workers, (iii) the western surrounding zone which is considered as areas with high skilled workers, (iv) the 2nd suburban belt and (v) the 3rd suburban belt. In that way, we capture not only the preference for the central areas, but also the preferences, if any, between those different zones. In addition, it is an attribute easily observable by the firm.

Last, we have introduced the price per square meter for different types of business premises. We have estimated semi-hedonic models were we have introduced location attributes as dependent variables, which are not present in our location choice model, in order to avoid high multicollinearity. We capture trade-offs between the positive attributes of a location, notably the accessibility or agglomeration effects and the price that a firm has to pay in order to be located to this area and enjoy these positive effects. For the Front Office services, we take the price for offices and for Back Office services we take the price for warehouses. It is expected to have negative influence ceteris paribus.

4.2.2. Location externalities

Location externalities seem to be the most undeniable determinant for a location choice of a firm (Hayter, 1997). They arise when firms use other establishments as a resource to their own productivity and from which a firm benefits. These location externalities or agglomeration economies can be divided in two different types: (i) the localisation economies or sectoral agglomeration economies and (ii) the urbanisation economies. As Marshal (1890) points out, the localisation economies or specialisation externalities emerge from the concentration of an economic sector to a specific geographical area. The proximity between firms in a specific industry can favour the labour market pooling, input/output sharing and knowledge spill overs (de Bok and Van Oort, 2011). It increases the performance of firms and reduces the risk for the implementation of new ones.

In empirical applications, sectoral agglomeration effects or MAR externalities (Glaeser et al., 1992) can be measured either by using the location quotient by economic sector or by the density of employment or firms (Beaudry and Schiffauerova, 2009). To measure sectoral agglomeration effect, after testing for all possible formulations, we have used the density of firms by zone and by sector, which gives the best, and the most consistent results. This agglomeration effect can also be a proxy for accessibility components that we mentioned before but not captured by our accessibility indicators. They are closely related to accessibility (de Bok and Van Oort, 2011; Melo et al., 2016) but from a broader firm-to-firm influence point of view without considering the influence of the infrastructure (de Bok, 2007) as it concerns the physical proximity for enterprises in the same industry.

The urbanisation or diversity externalities (Jacobs, 1969) are the result of the colocation of diverse economic sectors into a geographic area and of the increase in size of employment and population (Beaudry and Schiffauerova, 2009). Literature has not concluded if it has a positive or a negative influence on the location choice of a firm. It seems that its influence depends on the characteristics of each specific industry (Rosenthal and Strange, 2003). There are economic sectors, which value more the diversity and the density of a location while others are searching for more specialised locations. The urbanisation effects can be measured by the employment density, the Gini coefficient or the Hirschman-Herfindahl index (HHI) (Beaudry and Schiffauerova, 2009). We have opted for the HHI, modifying it as HHI'=1-HHI in order to have more intuitive results (Positive sign, positive influence of the local economic diversity).

4.2.3. Social environment

Other than accessibility and location externalities, we have also included the social environment of the location. Studies are not including social environment variables very often into the analysis. However, we are arguing that they can influence the location choice of a firm. Firms who offer high quality services and need face-to-face contact are expected to be located to areas where the revenues of the households are high (Elgar et al., 2009). Additionally, it is expected that firms should in general avoid areas with low-income households due to any possible social problems that can hurt the productivity of the firm. In addition, high-income households are attractive from a market potential point of view. Wealthy neighbourhoods can be potentially very attractive for certain sectors which value an area with good image, like for example the real estate agencies (Buczkowska and de Lapparent, 2014). We are taking into account the effect of the social environment by introducing into the model the percentages of the population of the zone belonging to the 1st quantile (Q1% - the poorest) and the 5th quantile (Q5% - the richest) of the revenue of the whole study area. If the sector relies on face-to-face interactions, it should be positive to the 5th quantile. On the contrary, the establishments should avoid areas with high 1st quantile. An area with high percentages of poor population can have possible negative local effects on the economic activity (Bouzouina, 2015).

4.2.4. Institutional factors

One of the latest advances of the location theory is the understanding that firms are making choices in an environment that is not static, because of government choices and real estate dynamics, the so-called institutional factors. In order to include in our model the effect of these factors we have included two variables. First, to account for the role of the macro-agent (public authorities, government) we have integrated a binary variable for the Zones of Economic Activity (Zones d'Activité Economique – ZAE). A ZAE is a designated geographic area of concentration of economic activity, organised and constructed by a public or private developers, which are renting or selling the land and the premises to enterprises willing to locate their businesses in these areas (Cerema, 2014). It is expected that this variable has a positive effect since in general these areas are located near transport axes and have created agglomeration or urbanisation effects.

5. Location choices of Front Office and Back Office business services – descriptive statistics

Before proceeding to location choice modelling we present a descriptive and cartographic presentation of the data in order to characterise our dataset and to provide an image of the dynamics across sectors.

In table III.1, we make a presentation of some key descriptive statistics regarding the two sectors. Most of the establishments of business services have a Front Office function. The Back Office sector has slightly bigger creation rate but the Front Office sector has a higher relocation rate. Regarding some preferences for proximity to transport infrastructures, we can see that Front Office establishments prefer better proximity to metro/tram stations while back Office establishments prefer proximity to motorways. Last, regarding the relocation distance, Front Office establishments are choosing locations that are closer to their previous one, in comparison to the Back Office establishments who can go further away.

***Table III.1** Key statistics of creations and relocations in Front Office and Back Office services*

		Front Office Services	Back Office Services
Total establishments		21702	9275
Creations	Establishments	9981	4130
	Creation rate	48%	49%
	Share near motorways	25%	37%
	Share near metro/ tram station	42%	30%
Migrations	Establishments	2940	989
	Migration rate	14%	12%
	Share near motorways	28%	39%
	Share near metro/ tram station	46%	27%
	Mean relocation distance	4,0 km	6,1 km

Source: INSEE SIRENE database 2011 - Authors' estimations.

***Figure III.2** Preference for accessibility to population of new and migrating firms*

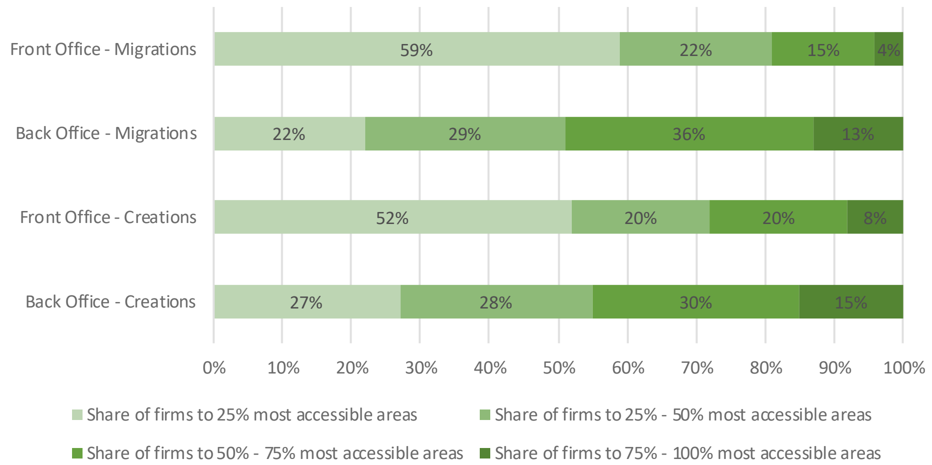
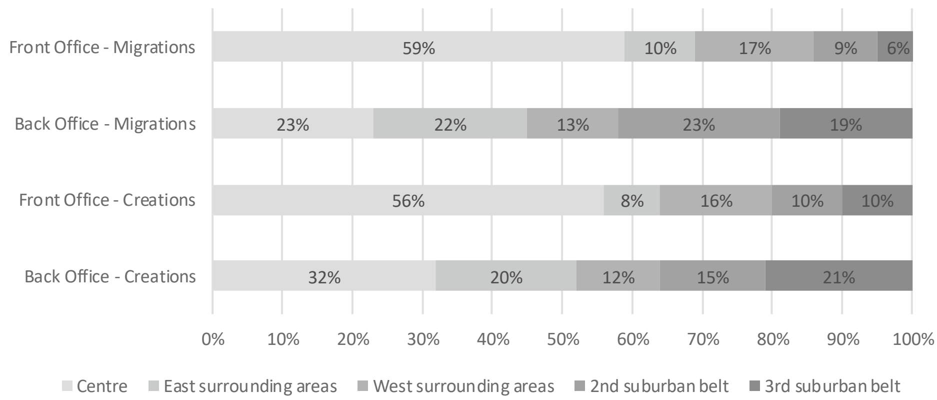
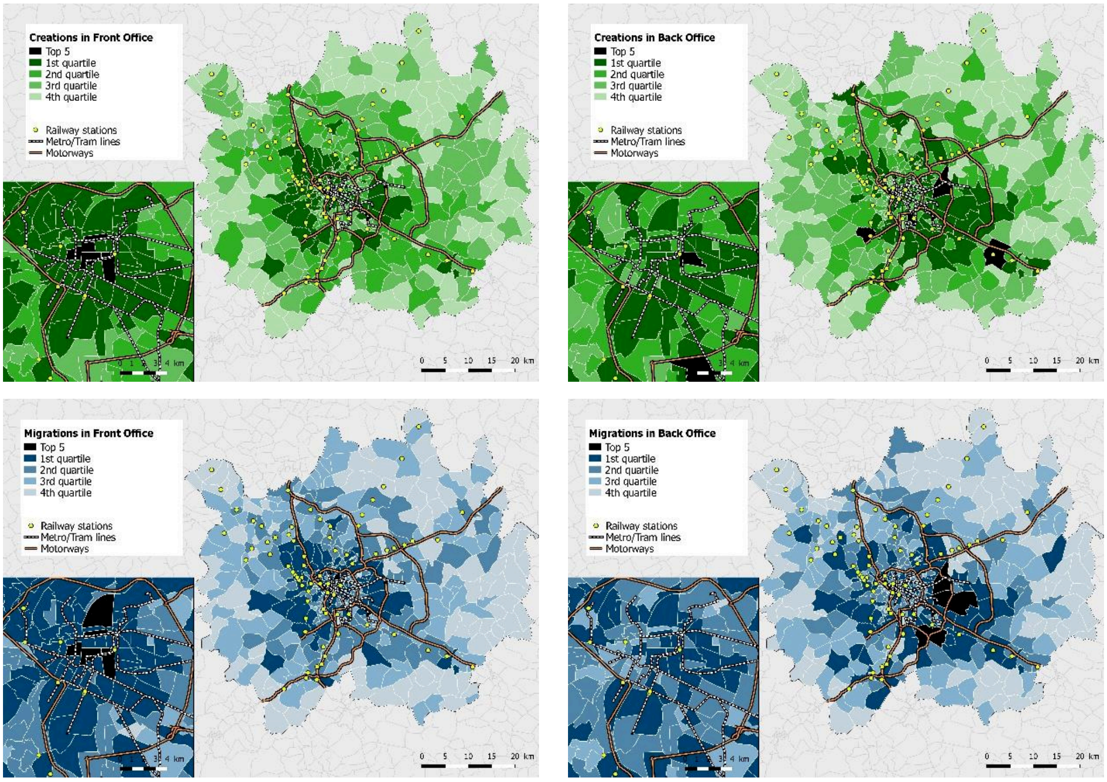


***Figure III.3** Location areas of new and migrating firms*



Source: INSEE SIRENE database 2011 - Authors' estimations.

In figures III.2 and III.3, we present an analysis of the preferences of Front Office and Back Office establishments for accessibility. We use two kinds of accessibility measures, the accessibility to population and the centrality. The first figure presents the distribution of establishments (creations and relocations) based on the accessibility of the selected zone (in quartiles). Front Office establishments are more sensitive to accessibility to population. More than 50% of created and relocated establishments of the Front Office sectors have chosen a zone between the 25% most accessible zones of our study area. On the contrary, Back Office establishments choose zones that are less accessible. Most of them choose zones that are at the third quartile of accessibility. Front Office activities have a particular preference for central areas, and seem to avoid the eastern surrounding areas. In contrast, Back Office services have no strong preference for any zone, but they seem to avoid the western surrounding areas.



Figures III.4-III.7 *Clockwise from the top left*
Figure III.4 *Location choices of new Front Office establishments*
Figure III.5 *Location choices of new Back Office establishments,*
Figure III.6 *Location choices of relocating Back Office establishments,*
Figure III.7 *Location choices of relocation Front Office establishments*

There is a clear difference between the two sectors. The Front Office establishments have high preference for central areas. This observation is consistent for both creations and relocations of Front Office establishments. On the contrary, Back Office establishments prefer mostly peripheral areas around the corridors of motorways, especially the eastern peripheral areas. This tendency is even more accentuated for the relocating establishments.

The spatial distribution of the location choices confirm the previous observations (figures III.4-III.7). Front Office establishments have high preference for the central areas, which decreases to the peripheral zones. This observation is consistent for both creations and relocations. On the contrary, Back Office establishments prefer mostly peripheral areas around the corridors of motorways, especially the eastern peripheral areas. This tendency is even more accentuated for relocating establishments.

6. Modelling results - Front Office, Back Office services: Diverging choices

The descriptive statistics analysis gave an idea of the dynamics and preferences for accessibility of the two subsectors. In this chapter, using a joint logit model for the Front Office and the Back Office establishments, we quantify the effect of the explanatory variables. We estimate a first model with the same variables for Front Office and Back Office (Model I) and an additional one integrating the variable "distance to last location" for the relocating establishments (Model II). The objective is to quantify the effect of this variable highlighted by the literature in firm migrations (de Bok and Van Oort, 2011; Elgar et al., 2009; Van Dijk and Pellenbarg, 2000). To compare the results across sectors and firm events, we have calculated mean point elasticities for the quantitative variables and mean pseudo-elasticities for the categorical ones (Washington et al., 2011). Last, we apply the asymptotic t-test (Ben-Akiva et al., 2015) between the two subsectors in order to test if the observed differences of the parameters are statistically significant.

Mean point elasticity	$E_k = \frac{\sum_{i=1}^{j}(1-P_{in})\,\beta_k\,X_{ink}}{j}$	*E_k: Mean point elasticity for quantitative variable k* *j: The number of establishments* *P_{in}: The probability for individual i to choose the location n* *β_k: The estimated parameter for k* *X_{ink}: The value of k for i at n*	(III.6)
Mean pseudo-elasticity	$E'_k = \frac{\sum_{i=1}^{j}\frac{P_{in}(i\mid X_{ink}=1)-P_{in}(i\mid X_{ink}=0)}{P_{in}(i\mid X_{ink}=0)}}{j}$	*E'_k: Mean pseudo-elasticity for categorical variable k* *P_{in}: The probability for individual i to choose the location n* *X_{ink}: The value of k for i at n which can take the values 0 and 1*	(III.7)
Asymptotic t test	$t_{k12} = \frac{\beta_{k1}-\beta_{k2}}{\sqrt{Var(\beta_{k1})+Var(\beta_{k2})-2Cov(\beta_{k1},\beta_{k2})}}$	*t_k: t value for variables k_1, k_2* *β_{k1}: The estimated parameter for k_1* *β_{k2}: The estimated parameter for k_2*	(III.8)

The significance of the variables varies between sectors and between firm events. This means that the economic sectors are making their location choices based on different criteria.

Front Office establishments appreciate accessibility to population (table III.2), all the variables of proximity to transport infrastructure are positive and significant and prefer central areas (all alternatives have a negative or not significant parameter). The premises' price has the expected negative and significant parameter, which characterize the trade-off between accessibility and land value. Last, the results for the other groups of variables have the expected parameters. Back Office establishments on the contrary, while they value the proximity to transport infrastructure, seem indifferent to accessibility to population (parameter non-significant) and prefer peripheral areas than central. However, the parameter for the premises' price is negative and significant meaning that there is a trade-off between the location attributes and the price. The parameters of the other variables have the expected signs. For both sectors, these results are consistent for creations and relocations, except for the accessibility to population for migrating establishments in the Front Office services, which we will discuss in the analysis of the elasticities.

The result of the estimation of the mean point elasticities shows the mean effect of a location attribute on the probability to choose a location, when we increase its value by 1%, ceteris paribus (table III.3). If the result is more than 1% (absolute value) it means that this variable is elastic (Washington et al., 2011). The result of the mean pseudo-elasticity shows the mean effect of a location attribute on the probability to choose a location, when it passes from zero to one. We can compare the results of elasticities and pseudo-elasticities between sectors and firm events, because we performed a joint estimation with the same variables. This is not applicable for the agglomeration effects (specialisation externalities) and the premises' price because they are sector-specific attributes.

For the Front Office services, the diversity of the location and the accessibility to population count the most for a location choice. There is a close relation between the diversity and accessibility. Central areas are more diverse while peripheral zones are more specialised (Manjón-Antolín and Arauzo-Carod, 2011). The elasticity for the diversity is 1.47% for the new establishments and 1.66%-1.73% depending on the model. The accessibility to population has a value of 1.24% for the creations and 1.37% for the relocations in model I. In model II it becomes negative, -0.35%, meaning that the establishment is willing to sacrifice accessibility in order to have a location choice to the previous one.

Table III.2 *Estimated parameters of the location choices of Front Office and Back Office services*

		Model I				Model II			
		Front Office		Back Office		Front Office		Back Office	
		Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.
Creations	Agglomeration	0.27	***	1.19	***	0.27	***	1.20	***
	Urbanisation	1.79	***	0.72	***	1.80	***	0.70	***
	Accessibility pop.	0.19	***	0.03	-	0.19	***	0.03	-
	Motorway	0.26	***	0.44	***	0.26	***	0.44	***
	Metro/Tramway	0.17	***	0.18	***	0.17	***	0.17	***
	Railway Station	0.26	***	0.12	***	0.26	***	0.12	***
	Centre (Reference)	-	-	-	-	-	-	-	-
	Eastern Areas	-0.60	***	0.15	**	-0.59	***	0.15	**
	Western Areas	0.02	-	0.19	**	0.02	-	0.19	**
	2nd Belt	-0.50	***	0.16	-	-0.50	***	0.16	-
	3rd Belt	-1.28	***	-0.55	***	-1.28	***	-0.56	***
	Q1 %	-1.01	***	0.10	-	-1.10	***	0.03	-
	Q5 %	3.26	***	-0.52	*	3.22	***	-0.54	*
	ZAE	0.59	***	0.70	***	0.59	***	0.70	***
	Premise's price	-0.32	***	-0.19	***	-0.32	***	-0.19	***
Migrations	Agglomeration	0.32	***	1.75	***	0.33	***	1.99	***
	Urbanisation	2.03	***	0.06	-	2.11	***	0.22	-
	Accessibility pop.	0.19	***	0.02	-	0.02	*	0.00	***
	Motorway	0.48	***	0.50	***	0.49	***	0.42	***
	Metro/Tramway	0.14	**	0.15	-	0.26	***	0.12	-
	Railway Station	0.39	***	-0.13	*	0.39	***	0.01	-
	Centre (Reference)	-	-	-	-	-	-	-	-
	Eastern Areas	-0.48	***	0.56	***	-0.13	-	0.46	***
	Western Areas	-0.01	-	0.33	*	0.00	-	0.17	-
	2nd Belt	-0.71	***	0.59	***	0.00	-	0.71	***
	3rd Belt	-1.85	***	-0.68	**	-0.02	-	0.38	-
	Q1 %	-1.60	***	-2.61	***	-1.80	***	-2.96	***
	Q5 %	3.02	***	0.00	-	2.71	***	0.81	-
	ZAE	0.68	***	0.91	***	0.72	***	0.79	***
	Premise's price*	-0.33	***	-0.38	***	-0.35	***	-0.53	***
	Distance last loc	-	-	-	-	-0.32	***	-0.27	***
Observations (segment creations)		9981		4130		9981		4130	
Observations (segment migrations)		2940		989		2940		989	
Observations (segments total)		12921		5119		12921		5119	
Observations (total)		18040				18040			
Alternatives		431				431			
Likelihood zero		-109433				-109433			
Log of Likelihood		-98324				-95544			
Adjusted ρ^2		0.102				0.127			

***significant at 99%, **significant at 95%, significant at 90%

Source: Authors' estimations.

Table III.3 *Estimated mean elasticities for Front Office and Back Office services*

		Model I				Model II			
		Front Office		Back Office		Front Office		Back Office	
		Elasticity	Signif.	Elasticity	Signif.	Elasticity	Signif.	Elasticity	Signif.
Creations	Agglomeration	0.86%	***	0.36%	***	0.86%	***	0.37%	***
	Urbanisation	1.47%	***	0.58%	***	1.47%	***	0.57%	***
	Accessibility pop.	1.24%	***	0.15%	-	1.24%	***	0.16%	-
	Motorway	30.21%	***	55.70%	***	30.23%	***	55.92%	***
	Metro/Tramway	18.92%	***	19.36%	***	18.55%	***	19.09%	***
	Railway Station	29.59%	***	13.05%	***	29.73%	***	12.94%	***
	Centre (Reference)	-	-	-	-	-	-	-	-
	Eastern Areas	-44.90%	***	16.41%	**	-44.66%	***	16.53%	**
	Western Areas	2.45%	-	21.26%	**	2.29%	-	21.25%	**
	2nd Belt	-39.25%	***	17.02%	-	-39.40%	***	17.00%	-
	3rd Belt	-72.14%	***	-42.55%	***	-72.16%	***	-42.82%	***
	Q1 %	-0.14%	***	0.02%	-	-0.15%	***	0.01%	-
	Q5 %	0.85%	***	-0.10%	*	0.84%	***	-0.11%	*
	ZAE	79.95%	***	101.49%	***	79.94%	***	101.20%	***
	Premise's price	-4.04%	***	-1.00%	***	-4.05%	***	-1.02%	***
Migrations	Agglomeration	1.26%	***	0.46%	***	1.30%	***	0.52%	***
	Urbanisation	1.66%	***	0.05%	-	1.73%	***	0.18%	-
	Accessibility pop.	1.37%	***	0.12%	-	-0.35%	*	-0.93%	***
	Motorway	61.92%	***	64.18%	***	63.49%	***	52.75%	***
	Metro/Tramway	15.45%	**	16.01%	-	29.90%	***	12.49%	-
	Railway Station	48.40%	***	-12.38%	*	48.27%	***	0.81%	-
	Centre (Reference)	-	-	-	-	-	-	-	-
	Eastern Areas	-38.22%	***	75.08%	***	-11.84%	-	58.50%	***
	Western Areas	-0.93%	-	39.30%	*	0.29%	-	18.57%	-
	2nd Belt	-50.94%	***	80.82%	***	0.00%	-	103.23%	***
	3rd Belt	-84.30%	***	-49.13%	**	-2.34%	-	46.60%	-
	Q1 %	-0.21%	***	-0.38%	***	-0.24%	***	-0.43%	***
	Q5 %	0.81%	***	0.00%	-	0.72%	***	0.17%	-
	ZAE	96.87%	***	147.43%	***	105.28%	***	120.00%	***
	Premise's price	-4.41%	***	-1.98%	***	-4.66%	***	-2.78%	***
	Distance last loc	-	-	-	-	-1.28%	***	-1.66%	***

***significant at 99%, **significant at 95%, significant at 90%
Source: Authors' estimations.

Migrating establishments search for agglomeration effects as well (elasticity 1.26%-1.30%) and they are sensible to distance from the last location (elasticity 1.28%). At the same time, Front Office establishments are very sensible to the premises' price with an elasticity over -4% for both creations and relocations. Regarding the proximity to transport infrastructures, we can see that the most important one for the Front Office services is the motorway. Its presence increases the probability to choose an area by 30.21% for the creations and 61.92%-63.49% for the relocations. However, an Economic Activity Zone is more important for a location choice of a Front Office establishment. Its presence increases the choice probability by 79.95% for the new establishments and 96.87%-105.28% for the relocating ones. For Back Office services, all the quantitative variables have elasticities less than 1%, except for the premises' price, which is -1% for the creations and -1.98% for the relocations. Even though Back Office establishments seem indifferent towards accessibility to population, they appreciate proximity to motorway, which has the highest pseudo-elasticity (55.70% and 64.18%-52.75 for creations and relocations respectively) between the transport infrastructures. As for the Front Office services, the presence of an Economic Activity Zone increases the probability to choose the zone by 101.49% for a new establishment and by 147.43%-120% for a relocating one. Last, migrating Back Office services are also sensible to distance to the previous location with an elasticity of 1.66%. Relocating establishments of both subsectors seem to optimise their location choices. They choose better locations in terms of accessibility, stronger location externalities to Economic Activity Zones to a better price.

When we compare the results between the Front Office services and the Back Office services, we can see a difference in preferences. On the one hand, Front Office establishments value more the central areas, the diversity, the accessibility to population of a zone and the proximity to a railway stations than the Back Office establishments. Additionally, Front Office firms are more sensible to the social environment of the zone. They avoid poor areas and prefer zones with high percentages of rich households. On the other hand, Back Office establishments have a higher preference for proximity to motorways, for peripheral areas and for Economic Activity Zones.

To analyse the statistical significance of observed difference between the two sectors, we applied the asymptotic t-test (Ben-Akiva et al., 2015). Usually this test is applied to the parameters of a model in order to verify that they are statistically different for zero. In our case, we are applying the test to verify that the parameters of the different sectors are statistically different between them. If the t value is greater than the critical value (in our case at 95%

confidence for 14 degrees of freedom for Model I and 15 degrees for the migrations of Model II) then the parameters are statistically different between the two sectors.

Table III.4 *Results of the t-test for the variables of the Front Office versus Back Office services*

	Variable	t value	Significance	More important for
Creations	Agglomeration	NA	-	-
	Urbanisation	-4.10	*	Front Office
	Accessibility pop.	-6.36	*	Front Office
	Motorway	4.14	*	Back Office
	Metro/Tramway	0.06	-	-
	Railway Station	-3.13	*	Front Office
	Centre (Reference)	NA	-	-
	Eastern Areas	9.07	*	Back Office
	Western Areas	1.80	*	Back Office
	2nd Belt	5.21	*	Back Office
	3rd Belt	4.55	*	Back Office
	Q1 %	-2.90	*	Back Office
	Q5 %	-10.68	*	Front Office
	ZAE	2.23	*	Back Office
	Premise's price	NA	-	-
Migrations	Agglomeration	NA	-	-
	Urbanisation	-3.83	*	Front Office
	Accessibility pop.	-3.43	*	Front Office
	Motorway	0.16	-	-
	Metro/Tramway	0.04	-	-
	Railway Station	-5.94	*	Front Office
	Centre (Reference)	NA	-	-
	Eastern Areas	6.31	*	Back Office
	Western Areas	1.77	*	Back Office
	2nd Belt	5.18	*	Back Office
	3rd Belt	3.60	*	Back Office
	Q1 %	1.26	-	-
	Q5 %	-4.40	*	Front Office
	ZAE	2.90	*	Back Office
	Premise's price	NA	-	-

* Significant at 95%.
Source: Authors estimations.

In table III.4, we present the results of the asymptotic t test. We can give an interpretation to the sign of the t-value. When the t-value is positive it means that parameter, in absolute values, is higher for the Back Office establishments. Otherwise, when it is negative, the absolute value of the parameter is higher for the Front Office establishments. Thus, depending on the sign and the significance we can conclude if the difference is statistically significant and for which sector this location attribute in more important. Diversity, accessibility to population, proximity to railway station and Q5% are more important for Front Office services. On the contrary, proximity to motorway (for the creations only), peripheral areas, Q1% (for the creations only) and Economic Activity Zones are more important for Back Office services. Last, proximity to metro/tramway stations, proximity to motorways (for the relocations only) and Q1% (for the relocations only) do not differ significantly between the Front Office and the Back Office services.

7. Conclusions and perspectives

Even though accessibility is considered as a key location factor for explaining the location choices of firms, not many works have analysed accessibility as a multidimensional concept. Even fewer have analysed any potential differences in terms of preferences for accessibility between different types of economic establishments. In this work, we examined the effect of accessibility of the location choices of the business sector, differentiating the establishments based on their functional characteristics. Most of the works on the location choices of economic establishments are analysing their behaviour based on the sectoral division of establishments. However, the literature points out that the functional categorisation tends to be more important that the sectorial (Duranton and Puga, 2005). Drawing from our analysis for Lyon, we observe that this categorisation defines the importance of the accessibility for the location choices of firms depending on their function. Our analysis highlights that each element of accessibility is valued differently from the different functions of the business services.

The functional characteristics of the economic establishments are reflected on their preferences for accessibility. The location choices of Front Office services, whose economic activity relies on daily interaction and information exchange (Dubé et al., 2016), follow ideally the traditional trade-off between premise's price and accessibility. Comparing to the Back Office services, the Front Office services have a stronger preference for central areas, where they can enjoy very good accessibility to population and where the location externalities are strong. On the contrary, Back Office services can offer their services by distance because direct proximity is less important for their economic activity. So, they prefer peripheral areas, at Economic Activity Zones, where they have easy access to motorways. The relocations of both subsectors prefer new locations close to the previous one, because of the importance of the local environment. They do not want to alter the relations with their clients, workers and collaborators.

This work can help the decision and policy makers to construct better policies, especially in transport. Investments on transport infrastructures or definitions of Economic Activity Zones can attract different types of economic establishments. An activity zone close to dense and populated area with high accessibility to population would attract different type of services than a zone at the periphery. Additionally, as this study showed, the attraction can be different for relocations and creations. Local policies should pay attention not to increase the economic activity locally at the expense of the nearby areas. Last, we support that future works should take into account the distinction that this article proposes based on the function of the business services sector because differences can be significant. A spatiotemporal analysis is needed to analyse if this functional division is always under way and if the preferences of the establishments evolve to a certain direction.

Chapter IV
Renters vs owners. How has the preference for accessibility evolved in residential location choices?

1. Introduction

Accessibility is an essential location attribute explaining residential location choices. Various works have shown its importance in different contexts (Baraklianos et al., 2018b; Eliasson, 2010; Lee et al., 2010). However, this importance might change and evolve over time. Changes that are internal to land-use and transport system (improvement or deterioration of the transport system, relocations of activities) or external (perception of individuals) can influence how accessibility is valued by households for their location choices (Kasraian et al., 2016; Portnov et al., 2011).

As there are new transport infrastructures and there is economic growth, accessibility is rising more and more. People can enjoy more activities at the same or a lower generalised cost (Handy and Niemeier, 1997). Consequently, as accessibility becomes more available, its importance for a residential location choice might decrease. This behaviour is predicted by the standard urban economics theory, the bid rent theory (Alonso, 1964; Mills, 1967; Muth, 1969). As accessibility increases, households seek larger properties to compensate for the reduction of the transport cost. In spite of the introduction of faster transport means, the stability of commuting around one hour – one hour and a half over the last century (Zondag and Pieters, 2005), seems to confirm this hypothesis. Transport improvements were capitalised into longer distance trips between residential and working locations (Axhausen, 2008). In the USA, the construction of motorways significantly contributed to the suburbanisation of population in the period 1950 to 1990 (Baum-Snow 2007). In Switzerland, the increase of accessibility led to a decrease of the importance of accessibility on the evolution of population during the period 1950-2000

(Portnov et al., 2011). In Lyon, the mean distance to work increased during the period 1985-2006 (Cabrera Delgado and Bonnel, 2016) mostly due to transport improvements. Other studies confirm that even investments on public transport infrastructures have contributed to the decentralisation of the population in Europe (Garcia-Lopez, 2012; Levinson, 2008).

During the last 20 years, despite the constant increase of accessibility of the urban areas, we have observed a revival of city centres across Europe (Buzar et al., 2007; Melia et al., 2018; Rérat, 2015) and the USA (Deka, 2018; Moos, 2013). Arguably, this re-urbanisation is mostly triggered by younger generations, the so-called millennials (roughly the generations born in the 80s and 90s (Deka, 2018)). Young households prefer to live in central locations with good accessibility, where public transport supply is satisfying, at the expense of spacious but car dependent suburbs (Deka, 2018; Melia et al., 2018). Various explanations have been given to this phenomenon (economic circumstances, anti-car culture, prolonged studies…) (Deka, 2018; Melia et al., 2018; Myers et al., 2019). Despite the various interpretations of this behaviour, a common denominator seems to be the preference for accessibility for residential location choices (Thomas et al., 2015).

A characteristic qualifying these two diverging residential choice behaviours is the type of choice (Haque et al.). On the one hand, households that make a long-term decision are less sensitive to accessibility. They seem to take advantage of the transport improvements to buy a residential unit in locations outside the highly accessible city centres. On the other hand, young households that make a short/medium-term choice are very sensitive to accessibility. They choose to rent an apartment in locations that offer high accessibility, despite the constant increase of accessibility.

A question that arises in this context is the extent to which these hypotheses are empirically confirmed. How different are the preferences for accessibility between renters and owners? More importantly, how have the preferences of these two groups evolved over time? In case of accessibility increase, do both two groups have the same reaction? If not, how do owners and renters adapt their location choices? To our knowledge, few works have analysed the residential location choices of these two groups together and even less the temporal evolution of their preferences. Concerning the importance of accessibility, consistent with the presented context, empirical works confirm its importance for the location choices of renters (Haque et al.; Inoa et al. 2015). Furthermore, this desire for accessibility seems to increase with time (Haque et al. ; Rezaei and Patterson, 2016), despite the increase of accessibility. However, either these works did not explicitly analyse separately the location choices of renters and owners (Rezaei and

Patterson, 2016) or they did not use data from distinct periods of time (Haque et al.). Analysing the temporal preferences of households is a very data-demanding task. Observations (e.g. household choices), estimations (e.g. generalised times) and socio-demographic data characterising the alternatives need to be available for multiple points in time. This might explain why works studying the temporal evolution of such preferences are rare.

In this paper, we aim to contribute to the understanding of the location choices of households bypassing some limits of previous works. Building on previous knowledge and taking advantage of the available data by means of previous works (Aissaoui, 2016; Cabrera-delgado, 2013; Homocianu, 2009), we analyse the evolution of the preferences of renters and owners over time, with a special focus on accessibility. For that, we rely on a discrete residential location choice model for the Lyon urban area in France. We use data for residential location choices drawn from the disaggregated census data of 1999, 2008 and 2013 and we estimate elasticities to measure the evolution of the sensitivity to accessibility and other location attributes. During this period, the city of Lyon has strongly increased its public transport supply by introducing new tramway lines and by extending the metro network while there was little investment in heavy car infrastructure (Bouzouina and Nicolas, 2015). In parallel, the urban area faced an important population and employment growth. From 1999 to 2013, the population increased by 22% while the employment progressed by 17%. These circumstances provide us with an interesting case study.

The rest of the paper is structured as followed. Section two (related empirical studies) presents the relevant empirical studies. Section three (study area) presents the study area and the evolution of the population and the transport infrastructure during the analysis period. In section four (method and data) we present the applied methodology and in section five (results and analysis) the results. In the last section (conclusions), we discuss the results and their implications along with the limits and the perspectives of our work.

2. Related empirical studies

Most of the works on residential location choice analyse a single period of time and focus on certain household groups, which are based on the socio-economic characteristics, such as revenue, size, age, car ownership, employment status, education level etc. What is less common is to make a more market oriented analysis using the housing occupation status as a choice determinant.

There are empirical works that analyse the differences between renters and owners. They converge on the idea that accessibility is more important for renters. This preference is related to the life stage of the household, which explains certain location decisions (Wilson et al., 2007). Specifically, Waddell (1996) analysed the residential location choices of households for Honolulu and confirmed that owners are less sensitive to accessibility to employment than renters. Plaut (2006) analysed the residential choice decisions in the USA of renters and owners in order to explain the mobility decision of dual-income households. He found that owners commute farther away. Inoa et al. (2015) studied the location choice decisions of households in Paris and found that renters appreciate better the accessibility of a location than owners. They relate this to the household's life cycle. A young household chooses a location with good accessibility in order to maximise all potential activities and employment. At later stages, when the work place is more stable and the household has children, a location choice at less accessible areas with other advantages is more likely. The authors mention that the location choices of the owners are more rational because they are more long-term choices, so they are more thoughtful.

Although the observation that renters have higher sensibility to accessibility than owners is not new, recent works have highlighted that this tendency has been reinforced after the 2000s. This trend is identified as a revitalisation process of the city centres due to the massive move of young households (Millsap, 2016). This tendency is highlighted in the works of Florida (2004) and Glaeser et al. (2001), suggesting that the reason behind this re-densification of the city centre is the presence of rich amenities in central - high accessible areas. In the few empirical works, analysing the temporal evolution of the households' location preferences the results are not conclusive. Rezaei and Patterson (2016), making a temporal analysis of residential location choices in Montreal between 1996 and 2006, found that households are becoming more sensitive to accessibility with time. However, the authors do not distinguish in their analysis between owners and renters for the accessibility variable, but the majority of the sample were renters (69%). We assume that this observation is due to the behaviour of renters. Furthermore, we do not know what the evolution of accessibility was like during that time. Any observed changes could be due to the accessibility improvements or deteriorations. A study for London (Haque et al.) analysed the temporal evolution of the residential choices of renters and owners, until 2002 and found a different result. The authors argue that owners became more sensitive to distance to city centre while renters left city centre because of high rents. However, both types of households without cars became more sensitive to the proximity to Public Transport. While these results seem surprising, this study presents some limits. First, the linear distance does not capture any possible differences of the accessibility throughout time. Second, the use of the same data to characterise alternatives for various observations in time, poses a methodological problem. We do not know the level of location attributes at the time the choice was made. Thus, parameters can be biased.

3. Study area

Our study focuses on the Lyon urban area in France. It has a population of 1.87 million and around 0.8 million jobs (INSEE, 2013). The urban area has a surface of about 3,3 thousand square km. It is the second largest metropolitan area of France after Paris in population and economic terms.

3.1. Evolution of the population

During the past 50 years, the urban area has increased its population by almost 50%. However, this evolution was not uniform. Like the majority of cities in Europe and the USA, Lyon has experienced the suburbanisation process. During the post-war era and until the 1990s, the population of the city centre decreased dramatically, by 18% in the period 1968 to 1990 while the population of the whole area increased by 23% (figure IV.1). From 1990 and afterwards, the tendency has changed and the population of centre started increasing. In the period 1990 to 2013, the population of centre increased by 22% surpassing the increase rate of the whole area, which was 20% during the same period.

From 1999 and onwards, the increase of the population of the centre was principally due to renters (figure IV.2). While the number of owners-households increased marginally, the number of renters-households presented an increase of 50%. On the contrary, the owners drove the growth of the other zones (2^{nd} belt and 3^{rd} belt) in the period 1999 to 2013 (for the aggregate zoning in five greater zones see figure IV.3). This tendency of return to the city centre is highlighted for Lyon in the literature by Aissaoui et al. (2015) and Rosales-Montano et al. (2015). This observation seems to be related to the attractivity of central areas by younger generations (Deka, 2018; Melia et al., 2018; Myers et al., 2019). The analysis of the housing construction permits (table IV.A1 – appendix) does not reveal any strong relation with the number of moved households.

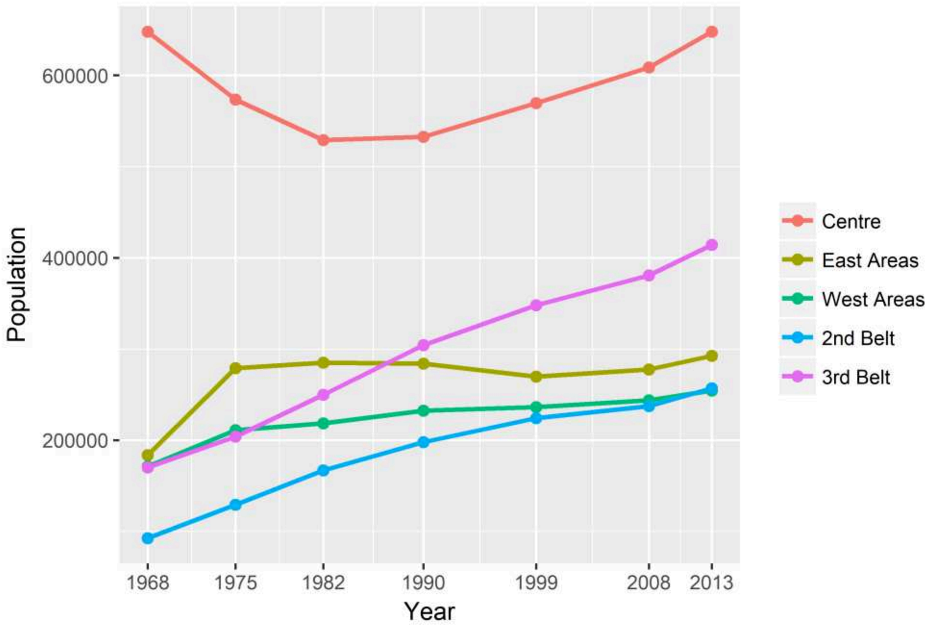


Figure IV.1 *Evolution of population from 1968 to 2013 by grater area*
Source: INSEE, authors' calculations

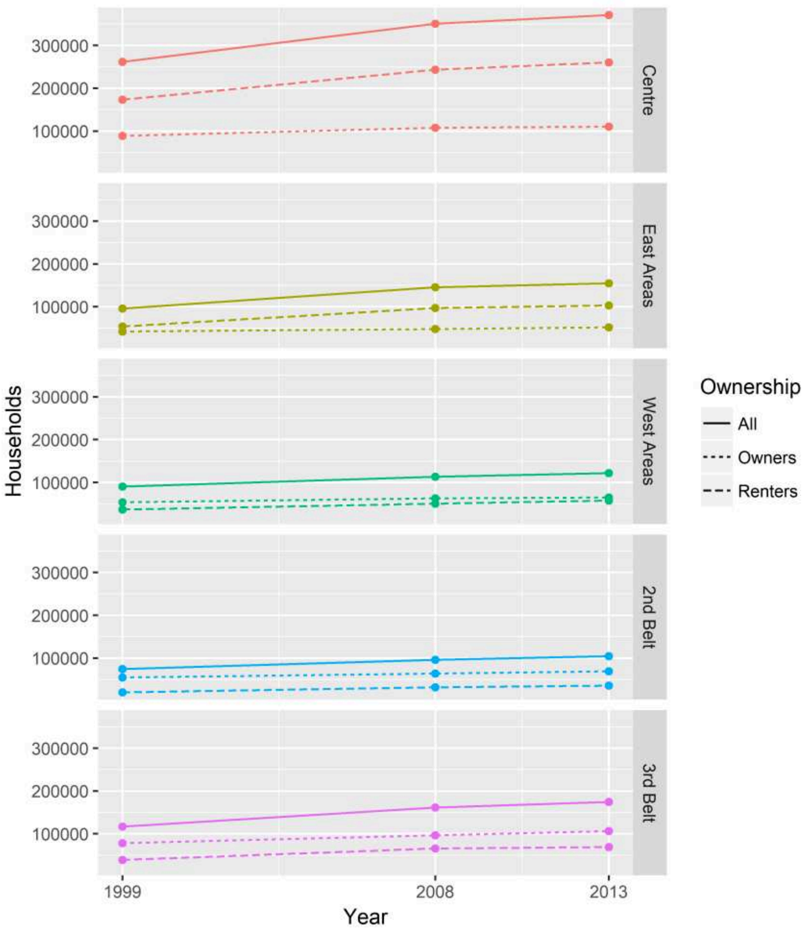


Figure IV.2 *Evolution of the number of households from 1999 to 2013 by aggregated analysis zone (figure IV.3) and housing ownership status*
Source: INSEE, authors' calculations

3.2 Evolution of public transport supply and accessibility

In the last 15 years, the public transport network of Lyon has been significantly improved. While the metro network changed marginally in the period 1999 to 2013 with three new stations, the investments on the tramway network were substantial. The first two tramway lines opened in 2000-2001, and three other new lines were inaugurated in 2006, 2009 and 2012, summing up to 58 kilometres of tramway network (figure IV.3). During the same period, there were not any significant car infrastructure investments. On the contrary, local authorities applied a policy aiming to decrease the importance of the car in the city of Lyon. Principally, they reallocated the urban space from car to public transport (tramlines, lanes for Buses with High Level of Service, junction priority to buses) or to other softer means of transport (bike lanes, pavement widening).

These investments together with the increase of employment (+17% in the analysis period 1999 to 2013) contributed significantly to the accessibility increase. Some areas benefited better from this improvement (table IV.1). The 3rd and the 2nd belt had the most important increase in accessibility to employment. Nevertheless, the 3rd belt is still by far the most deprived area in terms of accessibility. This favourable policy for public transport had a significant impact on the evolution of the modal shares in Lyon. Based on data from the Household Travel Surveys, car use decreased significantly for trips to work (from 67% in 1995 to 56% in 2015[1]), while the modal share of public transport progressed by 9% (from 12% to 21% respectively[8]).

Table IV.1 *Accessibility evolution by aggregate zoning zones*

	Centre	East surrounding areas	West surrounding areas	2nd urban belt	3rd urban belt	Total
Mean accessibility to employment (equation 6)						
1999	151,576	86,134	74,333	30,441	7,066	
2008	186,271	119,882	108,081	43,918	11,683	
2013	208,166	139,144	129,055	65,560	21,018	
Accessibility evolution						
1999 - 2008	+23%	+39%	+45%	+44%	+65%	+33%
2008 - 2013	+12%	+16%	+19%	+49%	+80%	+23%

Sources: INSEE, LAET, authors' calculations

[8] Own calculations from the household travel surveys data of the years 1995 and 2015.

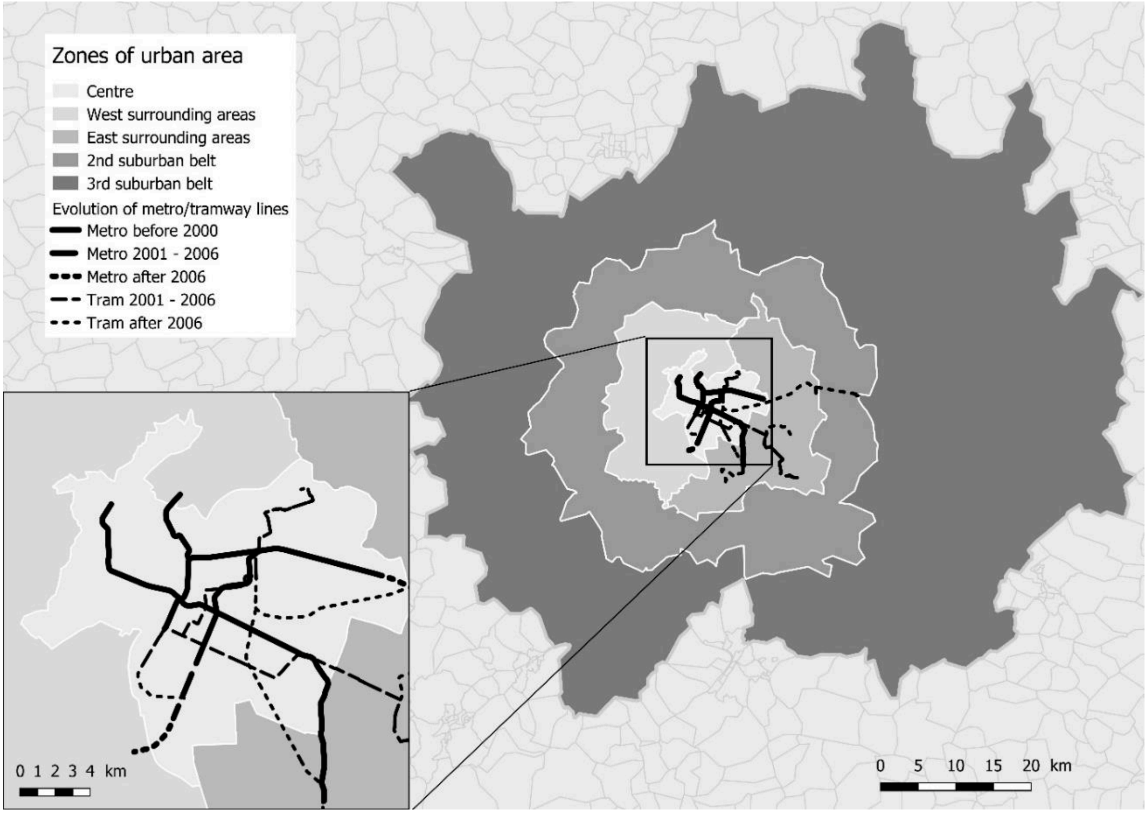


Figure. IV.3 *Aggregate urban area zoning and the metro/tramway network evolution*

4. Method and data

4.1. Residential location choice model

The modelling method used in this study is based on discrete choices (McFadden, 1977). In discrete choice modelling, the decision-maker selects the alternative from a choice set which maximises his utility (Schirmer et al., 2014). In residential location choice modelling, the decision-maker is the household, and the alternatives can be large zones, small neighbourhoods or even residential units. The individual utility is composed by a deterministic observable part and by a random term (equation IV.1).

$$U_{in} = V_{in} + \varepsilon_{in} \qquad \text{(IV.1)}$$

U_{in}: Utility of household n at location i
V_{in}: Deterministic part of the utility
ε_{in}: Error term

Under the assumption that the errors are independently and identically distributed - IID (McFadden, 1977), the probability of a household n making the choice i from a set of alternatives j, takes the logit form as is given by equation IV.2.

$$Pn(i) = \frac{e^{V_{in}}}{\sum_{j \in D_n} e^{V_{jn}}} \quad \text{(IV.2)}$$

V_{in}: Deterministic part of the utility of household n at i
V_{jn}: Deterministic part of the utility of alternatives j in D_n
D_n: Random choice set

In the Multinomial Logit (MNL) model, the deterministic/observable part of the utility depends on the attributes of the alternatives (zonal, dwelling etc.) and on the socio-demographic characteristics of the households. The utility function takes the form of equation IV.3.

$$V_{in} = \alpha X_i + \beta Z_{in} \quad \text{(IV.3)}$$

V_{in}: Deterministic part of the utility of household n at i
X_i: A vector of zonal attributes
Z_{in}: Interaction terms of socio-demographic characteristics of household n with the attributes of alternative i
α, β: Parameters to be estimated

A limit of the logit model is the assumption that the error terms are IID, which is unlikely in a spatial context (Ibeas et al., 2013). Other modelling structures like nested, cross-nested or mixed logit relax this hypothesis. However, these structures need an *a priori* assumption on the correlation structure and are difficult from an estimation point of view. Furthermore, they do not seem to improve empirically the modelling results for our case study (Aissaoui, 2016). Another option is to introduce spatially lagged terms into the deterministic part of the utility function (Alamá-Sabater et al., 2011; Rezaei and Patterson, 2016).

The problem of space arises because space is artificially divided into zones for modelling reasons. However, an attribute probably spills over into the contiguous zones. This spillover effect influences the utility of the observed choices. If it is not controlled, the model will estimate biased parameters. To control for this effect we extend the methodology applied to the firm location choice by Alamá-Sabater et al. (2011) into the residential location choice. We introduce spatial lag terms for variables that present important spatial autocorrelation into the deterministic part of the utility function (see Moran's 1 indices in table IV.3 for the spatial autocorrelation). Thus, equation IV.3 becomes equation IV.4. For the construction of the contiguity matrix, there are various methods. Alamá-Sabater et al. (2011) used a distance based matrix. In our case, we define the spatial lag terms in a simpler way. We use a binary contiguous zones (queen contiguity) binary matrix and calculate the spatial lag variables as the mean value of the attributes of these zones. A further complication of the spatial weights matrix has not improved the modelling results.

$$V_{in} = \alpha X_i + \beta Z_{in} + \delta(\alpha X_i' + \beta Z_{in}') \quad \text{(IV.4)}$$

X_i': A vector of spatially lagged variables
Z_{in}': A vector of spatially lagged variables for the interaction terms
δ: Spatial lag scale capturing the average influence of all the spatially lagged terms

The applied residential location choice model in our study is a Spatial MultiNomial Logit (SMNL) model and the alternatives are neighbourhoods. The study area is divided into 431 zones / neighbourhoods. This zoning is a census breakdown defined by the INSEE (Institut National de la Statistique et des Etudes Economiques - French National Institute for Statistics and Economic Research). These neighbourhoods have the advantage to be more detailed than the communes in the city centre and to respect a certain amount of homogeneity of the socio-demographic composition of the population (INSEE 2016). The estimation of a model with such a high number of alternatives is computationally difficult. When there is such a large number of alternatives, the parameters can be estimated using a random sample of alternatives D_n of the true choice set C_n and get consistent parameters (equation IV.2) (McFadden, 1977). We tested for various sample sizes of the D_n, up to fifty choices using various sampling strategies. We concluded that the best sample for the estimation is a random sample of seven random choices, the observed choice included, for every observed household choice (Aissaoui et al., 2015). This small random sample is sufficient and gives robust estimations with stabilized parameters.

In order to capture any potential systematic taste variations, we estimate market share models based on the housing occupancy status, distinguishing between renters and owners. The estimation is made jointly for the two groups. This distinction is important not only from a market point of view (owners are making a long-term choice while the renters are making a more short-term choice) but also from the point of view of the characteristics of those two groups (Haque et al., n.d.). Households that are owners tend to be richer, with more cars and with more household members, while households that are renters tend to be younger with less stable employment status.

For the estimation of the parameters, we need choice observations. The extraction of the observed residential choices was possible using the disaggregated census data provided by the INSEE, for the years 1999, 2008 and 2013. Using this database, we identify the households that moved in a certain neighbourhood during the last two years, the recently moved households. Consequently, we have 112,112 observations during 1998-1999, 102,920 observations in 2006-2008 and 120,623 during 2011-2013. The database contains information about the households like car ownership, number of individuals by household, the age of the household head and the status of the housing occupancy (owner/renter). The distribution by renters/owners and by area is presented in the table IV.2.

Table IV.2 *Distribution of recently moved households by aggregate zoning and housing occupancy status*

		Centre	East surrounding areas	West surrounding areas	2nd urban belt	3rd urban belt	Total
Recently moved households							
1999		55%	12%	12%	8%	14%	112,112
2008		55%	11%	11%	8%	15%	102,920
2013		53%	12%	12%	9%	14%	120,623
Housing occupation status of recently moved households							
1999	Renters	85%	80%	69%	57%	66%	78%
	Owners	15%	20%	31%	43%	34%	22%
2008	Renters	78%	65%	62%	52%	59%	69%
	Owners	22%	35%	38%	48%	41%	31%
2013	Renters	84%	74%	69%	58%	63%	75%
	Owners	16%	26%	31%	42%	37%	25%

Sources: INSEE, authors' calculations

4.2 Determinants of location choices

Studying the temporal evolution of residential location choice preferences does not only require the disaggregated data of the observed choices but also data on the attributes of the alternatives. This was possible through a combination of various databases. Ideally, data on location attributes must be for the beginning of the analysis period, to guarantee the exogeneity of these attributes. For example, for household choices made in the period 2006 to 2008, we need data characterising the alternatives for 2006. If we use the data of periods long after the observed choices, it is possible that the location attributes have changed, which would bias the parameter estimation. Nevertheless, it is very difficult to have the same historical data for the desired year. Because it was not always possible in our case, we had to do some compromises, while staying methodologically sound, which we describe in detail in this chapter. The variables included into the model are divided into three categories, the spatial amenities, the social environment and the market trade-off (Aissaoui 2016). Table IV.3 summarises the descriptive statistics of the zonal attributes. For a more detailed discussion of the variables, please refer to Baraklianos et al. (2019).

The first category captures self-segregation effects and preferences for social housing. The households have prefer a location that has a high concentration of the same revenue (Bouzouina, 2008). This set of variables can reveal a preference for endogenous amenities (Brueckner et al., 1999), which we cannot measure otherwise, like the quality of the services. To measure these self-segregation effects we introduce the percentage of households in a zone per revenue quantile (%REV3, %REV4, %REV5) as interaction terms with the same revenue quantile of the households (rev3, rev4, rev5). Regarding social housing, low revenue households should prefer zones with high social housing but high revenue households might avoid them. We introduce this variable as the percentage share of the social housing residential units of the zone (%HLM) interacted with all the five households' revenue quantiles (rev1, rev2, rev3, rev4, rev5). The data for the construction of the variables come from the databases of the INSEE[9] for the years 1999, 2006 and 2011. For these variables, we included spatial lag terms because we assume that the attributes of the contiguous zones can influence residential location choices.

The second category concerns the amenities that a household would value to have in close proximity (within the zone), like basic shopping services (bakery, supermarket, convenience store) primary schools and secondary schools in the zone. We expect that these variables would have a positive influence. The data for the estimation of these variables come from the SIRENE[10] database of the years 1999, 2006 and 2011. For these variables, we did not include spatial lag terms because they do not suffer from spatial autocorrelation (table IV.3). The inclusion of spatial lag terms for these variables did not increase the quality of the model.

The third category is the market trade-off between accessibility and housing price. Households make a trade-off between accessibility and price, as stated by the new urban economics theory (Alonso, 1964). We introduce the price as the zonal mean housing price per square metre. We estimate the housing prices from real estate transactions data from the periods 1998-1999, 2006-2007 and 2012-2013 (notary paid database - Perval). We introduce in all models the mean prices in euros2013 to account for the impact of the inflation. For the accessibility variable at location i, we use the potential accessibility to employment (equation IV.5) (Geurs and van Wee, 2004). This formulation captures accessibility changes due to transport network modifications (congestion, road network modifications, public transport improvement) or to employment increase. We use a composite generalised time, which combines generalised times by car and by public transport (equation IV.6) (for more details see in Bhat et al. (2002), Bhat et al. (1999) and Baraklianos et al. (2019)). A 4-step transport model estimates the times from zone to zone for each period by car and public transport (Cabrera-delgado, 2013) that permits us afterwards

[9] Historical data of the INSEE before 2006 not free. We used the one from 1999 because it was available at the laboratory

[10] SIRENE is a disaggregated database that contains all the companies in France and is provided by the INSEE (French National Institute of Statistics and Economic Studies). This database before 2016 was a paid database, so we used the most suitable data point for developing the residential location choice models

estimate the composite generalised times. The employment data come from the SIRENE database for the years 1999, 2006 and 2011. Despite the fact that these variables show important spatial autocorrelations, we did not include spatial lag terms. Accessibility is by construction spatially autocorrelated, meaning that the value of one zone already includes the effect of the contiguous zones. The housing price include the spatial correlation effect already so we do not have to control for it. Housing prices are influenced by locational characteristics shared by zones close to each other (Baumont, 2009).

$$A_i = \sum_j D_j e^{-\beta Tc_{ij}} \qquad \text{(IV.5)}$$

D_j: Number of jobs at destination j
Tc_{ij}: Composite generalised time from origin i to destination j

$$Tc_{ij} = \left(\frac{Tcar_{ij}}{1 + \frac{Tcar_{ij}}{Tpt_{ij}}}\right) \qquad \text{(IV.6)}$$

Tc_{ij}: Composite generalised time from i to j
$Tcar_{ij}$: Generalised time for car
Tpt_{ij}: Generalised time for public transport

Table IV.3 *Descriptive statistics and spatial autocorrelation of the zonal attributes*

Variable description	Variable	1999			2006			2011			Spatial Lag term
		Mean	SD	Moran's I	Mean	SD	Moran's I	Mean	SD	Moran's I	
Social environment											
Social housing share of households (%)	%HLM	0.11	0.18	0.43	0.12	0.17	0.39	0.12	0.16	0.42	Yes
Share of the 3rd quantile of revenue (%)	%REV3	0.23	0.04	0.30	0.25	0.07	0.37	0.22	0.04	0.41	Yes
Share of the 4th quantile of revenue (%)	%REV4	0.20	0.04	0.30	0.19	0.06	0.40	0.21	0.05	0.50	Yes
Share of the 5th quantile of revenue (%)	%REV5	0.21	0.10	0.51	0.19	0.12	0.54	0.22	0.10	0.52	Yes
Spatial amenities											
Proximity to basic shopping service (number) (0,1)	Prox Basic Serv	0.84	0.34	0.18	0.85	0.35	0.24	0.86	0.35	0.20	No
Primary schools (number)	Prox Pr. Schools	3.79	3.52	0.11	3.79	3.43	0.12	4.1	3.67	0.12	No
Secondary schools (number)	Prox Sec. Schools	0.54	0.97	0.01	0.54	0.92	0.00	0.69	1.35	-0.01	No
Market trade-off											
Mean zonal housing price (€ 2013/m^2)	Housing price	951	268	0.72	2,753	535	0.73	2,503	487	0.63	No
Accessibility to employment	Acc. Emp.	51,333	58,228	0.91	68,482	70,817	0.97	84,018	76,360	0.96	No

Sources: INSEE, INSEE-SIRENE, PERVAL, LAET, authors' calculations

5. Results and analysis

We discuss the results of the models before presenting the elasticities analysis. For each analysis period, we estimate all the models with the same independent variables distinguishing between renters and owners. Table IV.4 summarises the parameters.

The estimated parameters are significant and mainly have the expected signs. Accessibility is positive and significant for all periods and for both renters and owners. However, the parameter is higher for renters showing that they are more sensitive to accessibility. Concerning the other variables, the signs do not vary between the different periods, except for the interaction term %HLM*rev2 (social housing - 2nd revenue quantile household) for owners, which is not significant in 2008 and changes sign in 2013. This change shows a tendency that the presence of social housing decreases its importance for this revenue quantile.

Table IV.4. *Estimated parameters and significance of the models*
Source: Authors' calculations

	1999		2008		2013	
	Renters	**Owners**	**Renters**	**Owners**	**Renters**	**Owners**
Social environment						
%HLM*rev1	1.21**	0.18**	0.80**	0.35**	1.37**	0.15**
%HLM*rev2	0.49**	-0.62**	0.15**	-0.43**	0.11**	-0.98**
%HLM*rev3	-0.24**	-0.79**	-0.50**	-1.07**	-0.70**	-1.36**
%HLM*rev4	-0.10**	-0.64**	-0.55**	0.13	-0.63**	0.79**
%HLM*rev5	-1.28**	-1.70**	-1.91**	-1.39**	-1.18**	-0.76**
%REV3*rev3	4.66**	3.97**	5.22**	8.38**	5.51**	5.16**
%REV4*rev4	5.04**	5.96**	5.07**	9.83**	6.66**	9.03**
%REV5*rev5	2.77**	2.20**	2.38**	3.67**	1.51**	3.36**
Spatial amenities						
Prox Basic Serv	1.19**	0.74**	1.45**	1.02**	1.42**	1.06**
Prox Pr. Schools	0.10**	0.09**	0.09**	0.10**	0.07**	0.08**
Prox Sec. Schools	0.06**	0.02**	0.07**	0.08**	0.10**	0.09**
Market trade-off						
Housing price	-0.69**	-0.10**	-0.11**	-0.13**	-0.18**	-0.33**
Acc. Emp.	0.64**	0.43**	0.55**	0.26**	0.53**	0.21**
Spatial lag parameter (δ)	0.64**	0.64**	0.61**	0.27**	0.31**	0.14**
Observations (Total)	112,112		102,920		120,623	
Observations (renters/owners)	77,554	22,376	63,315	28,069	76,792	26,141
Log likelihood zero	-217,086		-199,400		-233,525	
Log likelihood (LL)	-159,666		-147,507		-175,905	
Rho-squared	0.264		0.260		0.247	

** significant at 95%

The spatial scale parameter is significant and has a value under 1. This means that the attributes of the contiguous zones have an effect on the utility of the chosen alternative, but their influence is smaller than the effect of the attribute of the chosen zone (Alamá-Sabater et al., 2011). This result is consistent. The value of the spatial scale parameter decreases between periods, which shows that the effect of the neighbouring zones is decreasing over time.

Due to scale differences between models estimated for different years, we cannot directly compare the parameters. All the variables, except for the proximity to basic services, are continuous, so we can calculate mean point elasticities (equation IV.7), which are directly comparable. Elasticities are also more convenient in terms of interpretation. The mean point elasticity shows the mean effect on the choice probability of an increase of 1% of this variable *ceteris paribus* (Washington et al., 2011).

Mean point elasticity	$E_k = \frac{\sum_{i=1}^{I}(1 - P_{in})\, \beta_k\, X_{ink}}{I}$	E_k: *Mean elasticity for variable k* *I: Number of households* P_{in}: *Probability of household i choosing the location n* β_k: *Estimated parameter for k* X_{ink}: *Value of k for i at n*	(IV.7)

In figure IV.4, we present the evolution of the mean point elasticities of the market trade-off variables (accessibility and housing price). The evolution of the other variables can be found in the appendix. Analysing the effect of accessibility alone is not consistent. This is why we focus on both market trade-off variables. Accessibility is constantly more important for renters, given that the elasticity values of renters are systematically higher than the values of owners. Our results are consistent with previous studies (Inoa et al., 2015; Plaut, 2006; Waddell, 1996).

The evolution of the elasticities of accessibility and housing price are very different between renters and owners. For renters, the elasticity of accessibility increases during the analysis period, meaning that renters become more sensitive, despite the strong increase of the accessibility levels. The elasticity increases almost linearly, from 1.71% in 1999 to 1.90% in 2008 reaching at 2.32% in 2013. The owners' elasticity for accessibility decreases during the analysis period, from 1.02% in 1999 to 0.67% in 2008 and to 0.57% in 2013. At the same time, the evolution of the elasticity of the housing price follows the same tendency, as expected. Renters become less sensitive while owners present the opposite tendency. For renters, the evolution is not linear but the elasticity becomes less negative. For owners, the decrease of elasticity is sharper and almost linear. The evolution of the elasticity for housing price is almost parallel to the elasticity for accessibility, proving the almost perfect trade-off for owners.

It seems that the overall increase of accessibility had different effect on household preferences depending on their housing occupancy status. The improvement of accessibility led the owners to be less sensitive to accessibility, confirming the bid-rent theory on the adjustment of location choices after an accessibility increase. The increasing sensitivity to the housing price (more negative elasticities) is consistent with this behaviour. On the contrary, renters do not seem to alter their location choice behaviour and they choose areas that increase their accessibility over time. We can see that the improvement of accessibility is capitalised into an increase of the elasticity for renters. The evolution of the elasticity of the housing price shows that renters are ready to pay more in order to access to high accessibility locations.

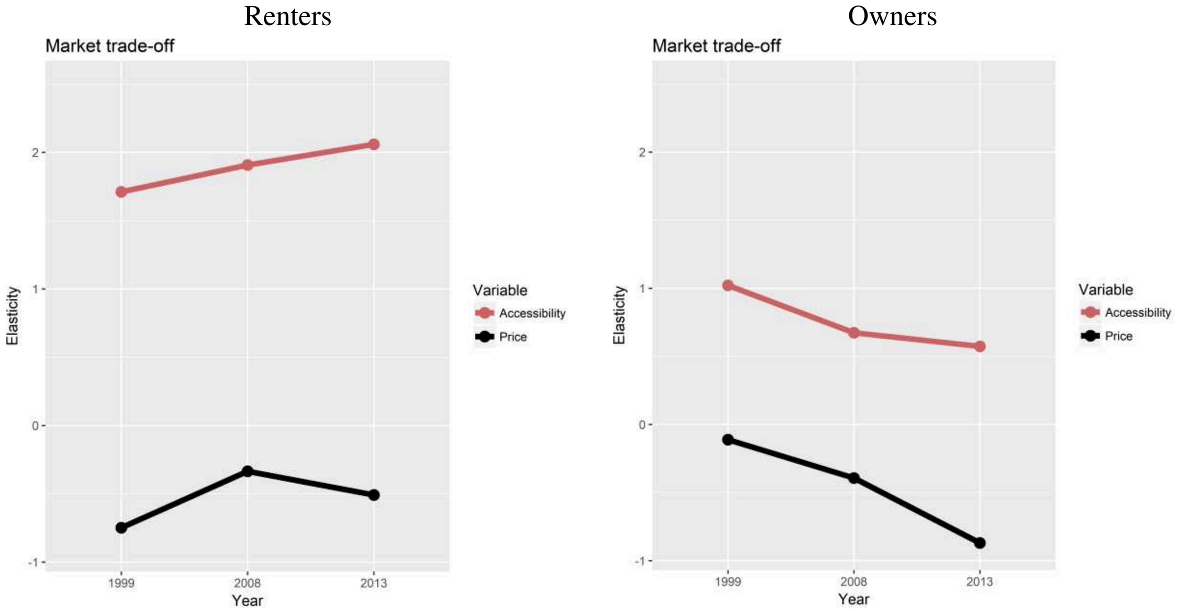


Figure IV.4 *Evolution of elasticities by variable group and by renters/owners*
Source: Authors' estimations

While the bid-rent theory can explain the behaviour of owners, the interpretation of renters' choice behaviour is not straightforward. First, this reaction can be related to the urban amenities theory developed by Glaeser et al. (2001). The city centre offers diverse opportunities for participation and remains attractive throughout the time. Possibly for renters, proximity and participation to these activities is highly valued because of their socio-economic characteristics (Melia et al., 2018). Second, there can be an evolution of the lifestyle of these households. Renters are more likely to be young adults forming households of one or two individuals without children. As recent literature points out, millennials create families at an older age than the previous generations (Deka, 2018; Melia et al., 2018; Oakil et al., 2016). This fact has an impact on their lifestyle choices and accessibility is an important location characteristic of their residential choices (Thomas et al., 2015).

We tried to find some relations between the observed location choice behaviour and the characteristics of households. The socio-demographic characteristics of the moved households are amazingly stable between the three analysis periods. The distribution of age and education level of the household's head, the number of household's members and the surface of the residential unit are the same for all periods. Nevertheless, a significant difference is found for car ownership. We observe that renters own fewer and fewer cars (table IV.5). The share of the recently moved renters without a car has risen from 35% in 1999 to 41% in 2013, while at the same time the share of renters with one car fell from 49% in 1999 to 44% in 2013. Various studies have shown that the new generations are less likely to own a car or hold a driver's licence than the previous ones (Klein and Smart, 2017; Melia et al., 2018; Oakil et al., 2016). In Lyon, we observe the same tendency (Vincent-Geslin et al., 2017) in line with our findings. This behaviour can explain why the elasticity of renters for accessibility increases. Because they do not own a car, deliberately or not.

Table IV.5. *Car ownership based on the housing occupancy status of the recently moved households*

Car ownership	No car		One Car		Two or more cars	
Housing occupancy status of recently moved households	Owner	Renter	Owner	Renter	Owner	Renter
1999	7%	35%	48%	49%	45%	16%
2005	6%	37%	46%	48%	48%	15%
2013	8%	41%	48%	44%	45%	15%

Sources: INSEE, authors' calculations

6. Conclusions

Many studies highlight the importance of accessibility for a residential location choice (Baraklianos et al., 2018; Eliasson, 2010; Lee et al., 2010). However, it is not clear in the literature how the preference for accessibility evolves over time. As accessibility levels rise by virtue of transport investments and economic growth, households are likely to adapt their preferences for accessibility. Based on the theory of urban economics (Alonso, 1964) and the observed urban sprawl of the post-war era (Baum-Snow, 2007), we can conclude that the increase of accessibility drives households to own residential units on peripheral areas leading to a decreasing preference for accessibility. On the other hand, during the last 20 years, the city centres have regained their lost attractivity (Cheshire, 2006). This regeneration of the city centres is driven by younger generations, the so called millennials, who prefer to rent a residential unit in highly accessible areas (Deka, 2018; Melia et al., 2018; Myers et al., 2019). In this changing environment, there are not many empirical works, which could help us understand how the preferences of the households for accessibility evolve over time.

A characteristic that seems to define the preference for accessibility is the housing occupation status (Haque et al.). The renters are more sensitive to accessibility (Inoa et al., 2015; Plaut, 2006; Waddell, 1996). Nevertheless, the few empirical works on the evolution of households for accessibility do not give us a clear picture. In this article, taking advantage of the available cross-sectional data, we relied on discrete choice models and analysed the evolution of the preferences for accessibility. Distinguishing between renters and owners, we estimated elasticities for accessibility to employment by car and Public Transport for the urban area of Lyon, for three periods 1999, 2008, and 2013.

The results for the urban area of Lyon confirm the findings of the empirical literature about the sensitivity of the renters. The evolution of the preferences shows that, despite the fact that renters already had higher preference for accessibility at the first analysis period (1999), the difference between owners and renters increases over time. Owners show a decreasing preference for accessibility, while renters the exact opposite. For owners this means that they have a more "rational"[11] location choice behaviour (Inoa et al., 2015). The increase of accessibility led to a decreased preference for accessibility, consistent with the urban economics theory. On the contrary, renters, which make a short-term decision, can be less rational. They present an increasing preference for accessibility. This choice behaviour is related to their socio-demographic profile, which explains why accessibility is very important for their residential location choices (Melia et al., 2018). They are more likely to be younger, at the beginning of their career and more importantly without children (Plaut, 2006). It seems that there is a relation between the car ownership levels and the preferences for accessibility, without being able to define a clear causality.

For policy makers, our results suggest that future transport planning and land-use policies should be more integrated. Changes of the public transport supply impact the choices of the households and thus the urban form (Garcia-Lopez, 2012; Levinson, 2008). Different urban dynamics can be observed depending on the modification of such choices. Planners should anticipate those changes and channel the household choices through housing supply in order to avoid the exclusion of certain households. As our results suggest, some households can be constrained to choose a location at the expense of other choices. Thus, public authorities should pay attention to offer residential units in accessible areas for sensible socio-economic groups.

These policy directions are even more opportune for the present and near future. Today, city centres face important challenges with the development of housing sharing schemes like Airbnb. Tourists and renters seem to have the same preference for accessibility, but the profitability of these solutions drive prices up leading to shortage of housing offer for local residents (Deboosere et al., 2019). In the decades to come, a radical change of the accessibility levels can be due to new mobility solutions like Mobility as a Service or automated vehicles (Meyer et al., 2017). Our results are in line with Milakis et al. (2018), who suggest that this

[11] Regarding the observed evolution of the trade-off between accessibility and housing price

evolution can cause at the same time more suburbanisation and increased density of the central areas. We draw the attention of the public authorities to these two ongoing processes that can potentially lead to polarisation and social exclusion of some social groups.

From a scientific point of view, this work enhances the knowledge around residential location preferences for accessibility with an empirical study, which can contribute to the understanding of the urban dynamics. The temporal dynamics depend on the characteristics of the decision makers and urban modellers should be aware of these. Methods combining different cross-sectional data can capture these temporal variations of the preferences and help in estimating better predictive models (Rezaei and Patterson, 2016).

While our study gives some interesting insights on the subject, we could not overcome some limits. The analysed period is fourteen years and the time steps are not equal. A longer analysis period and fixed time steps could be more appropriate. However, the use of disaggregated and historical data poses many restrictions. Therefore, the choice of the analysed period and of the time steps is highly dependent on the availability of the data. Furthermore, the temporal analysis of other choices conditioned by residential location choices, like the car ownership, can be very interesting. Those choices are interdependent and a more profound understanding can help the construction of better transport and land-use policies.

7. Appendix

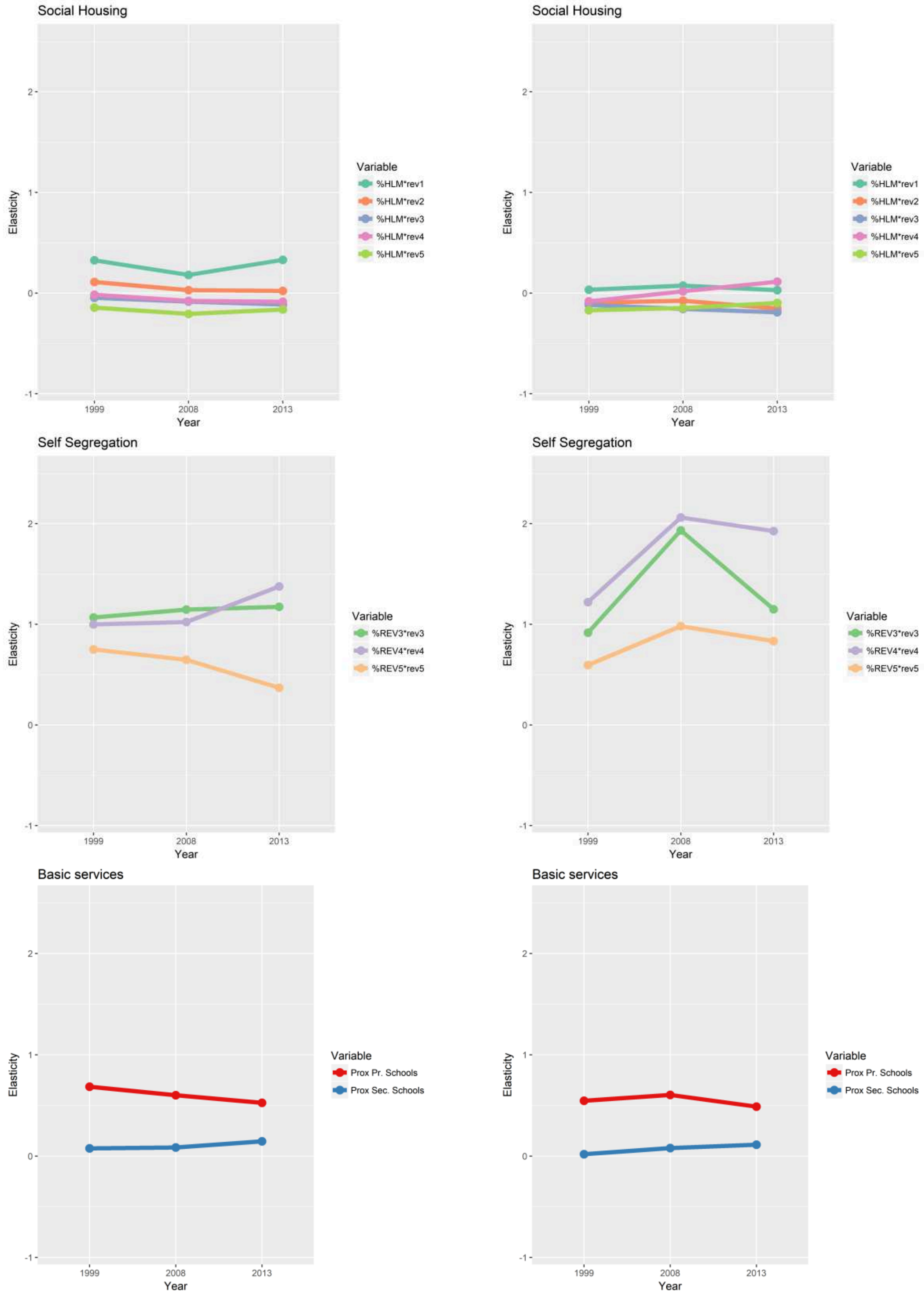


Figure IV.A1. *Evolution of elasticities by variable group and by renters/owners*
Source: Authors' estimations

Table I.VA1. *Distribution of authorised surface for housing by aggregated zone*

Zone	Centre	East surrounding areas	West surrounding areas	2nd urban belt	3rd urban belt	Total (m^2)
Housing construction permits 2006-2008	56%	2%	11%	1%	30%	77,352
Housing construction permits 20011-2013	39%	24%	5%	21%	11%	68,753

Source: Citadel database, Ministry of sustainable development

Conclusion

Population and employment in cities have always been in constant interaction with the transport system. A transport investment changes the attractiveness of any given location for households and firms. At the same time, the increase of population and economic activity boosts the need for transport investments. The Land-Use Transport Interaction (LUTI) models were born from the need to connect these two components in order to better understand and model their interactions. LUTI models aim to represent, simulate and forecast these kinds of interactions that arise in an urban environment, in the objective to help policy-makers to take informed decisions for future policies or projects.

Accessibility is a key construction element of LUTI models. As LUTI models consist of two main models, land use and transport, accessibility assures the interaction between them. Accessibility has this role because it translates any transport or land-use change into a quantifiable and measurable amount that can be integrated into the LUTI modelling chain.

However, accessibility is an intangible concept that we need to model. At the beginning, this modelling procedure was largely empirical with no theoretical framework (Geurs and van Wee, 2004). The theoretical bases came later, after accessibility proved that it is a valuable tool for analysing land-use and transport policies. Accessibility is theoretically constructed out of three components, (i) the land-use component, (ii) the transport component and (iii) the individual component (Geurs and van Wee, 2004). These components evolve in space and time. Empirically, the available methods to model accessibility can be classified in three groups in respect to the accessibility components, (i) the transport-based measures, (ii) the opportunity-based measures and (iii) the utility-based measures. However, each approach has its own pros and cons.

The modelling procedure implicates methodological choices to simplify a reality complex by nature. However, these choices are not neutral. They can influence the obtained results and the drawn conclusions. This is why the application of accessibility measures is considered to be a challenging process. LUTI models are no exception to the fundamental challenges of accessibility. Despite the rich literature on this subject, there are still a lot of LUTI–accessibility unanswered questions.

Between the different modules of LUTI models, this thesis focused on the location choice models of households and firms. Their interest goes beyond the LUTI models. Any integrated land-use transport policy must consider the locations of households and economic activities as the two anchor points of daily mobility. This thesis examined the place and the importance of accessibility in the location choice models of households and firms, analysing the effect of various methodological choices. Specifically, three main challenges were identified with relation to the integration of accessibility into location choice models:

(i) How does the inclusion of different components in the measurement method influence the location choice modelling results?

(ii) What are the implications of individual taste variation on location choice modelling?

(iii) What is the temporal stability of location choice preferences?

In the context of location choice models, the existing literature used various accessibility measures. One can either introduce a complex, comprehensive and complete indicator of accessibility, which covers all of the accessibility components, or introduce various simple accessibility measures, each covering one or more accessibility components. Within the plurality of existing approaches, we do not know the effect of methodological choices on the results of location choice models. Usually, studies use the measure that gives the best statistical results. Nevertheless, if we are not able to estimate complex enough measures that are theoretically better, would the result still be consistent? On the contrary, do complex measures confirm empirically their theoretical superiority? Does the inclusion of the accessibility components improve the obtained results?

Another important aspect regarding the inclusion of accessibility measures in location choice models is the integration of the individual component (Wee, 2015). Different households and firms value accessibility differently. In addition, as accessibility is a multicomponent measure, different accessibility components can have different influence on different individuals. A family household could value accessibility differently than a single-worker household. A brand-new firm in the business services sector could value differently proximity to a motorway than a relocating logistics company. New technologies transform how firms organise their functions spatially (Anas et al., 1998). Intuitively, those systematic differences exist but, in the framework of location choice models, are those differences significant? Do we need to develop distinct market share models? Do we need to use different types of accessibilities for different socio-economic groups?

Lastly, a dimension not thoroughly analysed in the literature of the location choice is the temporal evolution of the preferences. As the economic and technological contexts evolve, owing to advances and innovations, it is possible that the preferences of households and firms evolve as well. For households, literature indicates that newer generations show a lower preference to using their cars as a main mode of transport and use public transport more (Wee, 2015). Such changes that have already been studied in the framework of individual mobility might have important implications on location choices. As mobility habits and residential choices are interdependent, such evolutions should have great implications on the location choice preferences. There are already some indications that central areas become more and more attractive thanks to renters and younger generations, (Melia et al., 2018) but few works analysing such temporal dynamics exist. How does preference for accessibility evolve over time? Which households drive those changes?

Those three identified changes resulted into four research papers. First, **Paper 1 (chapter I)** examined the impact of the accessibility measurement method on the results of the residential location choice model. The knowledge obtained from this chapter permitted the appropriation of the accessibility measurement methods and the control for potential biases related to the modelling process. Building on this knowledge, **Paper 2 and Paper 3** analysed the individual dimension of accessibility focusing on firms. **Paper 2 (chapter 2)**, examined the influence of accessibility comparing creations and relocations of various sectors. **Paper 3 (chapter 3)** took a step forward and examined the differences within one of the economic sectors, the business services sector. Lastly, even if the previous chapters demonstrated the importance of accessibility, time might affect this preference. With this objective, **Paper 4 (chapter 4)** permitted the temporal analysis of the preferences for accessibility focusing on the differences between renters and owners for a residential location choice.

PhD thesis contributions

Does the accessibility measure influence the results of the residential location choice model?

Some empirical works have difficulties to find a significant relation between accessibility and residential location choice (Blijie, 2005; Lee et al., 2010). These studies have questioned the importance of accessibility. They suggest that other location factors like social environment, neighbourhood amenities and housing unit characteristics are more important in explaining residential choices (Blijie, 2005; Chen et al., 2008; Sener et al., 2011; Zondag and Pieters, 2005). Others consider accessibility to be essential in the estimation of residential location choice models (Eliasson, 2010; Srour et al., 2002). They suggest that the inclusion of accessibility in residential location choice models is important, regardless of the studied urban form (Lee et al., 2010).

In residential location choice models, a great variety of accessibility measures is applied (Schirmer et al., 2014). These measures vary from simple to complex ones. Research works include all types of accessibility measures (transport-based, opportunity-based, utility-based) depending on the estimation ability. Given the complexity of accessibility, translating its influence in a residential location choice context can be a challenging task.

The aforementioned controversy on the importance of the accessibility might be a result of the absence of a systematic comparison between measurement methods. The measurement method could affect the modelling results. In this context, the objective of **Chapter I** was twofold. First, to analyse the importance of accessibility as an explanatory factor even with simple accessibility measures and, second, to assess the benefit of using more sophisticated measures. To respond to these objectives, the empirical study relied on a residential location choice model developed for the urban area of Lyon. The estimation was based on observations from the period of 2006-2008. The explanatory variables characterise local spatial amenities, the social environment and the trade-off between accessibility and housing price. Various models were estimated, altering only the accessibility measure, while keeping all the other variables of the model the same.

Regarding the accessibility measures, the work focused on opportunity-based measures. We examined the results of five different accessibility measures: (i) proximity to transport infrastructure, (ii) linear distance to centre, (iii) generalised time to centre, (iv) cumulative opportunities to employment and (v) potential/gravity accessibility to employment. For the last indicator, an appropriate method is applied in order to aggregate between the two main transport modes, public transport and the car. The aggregation is based on a composite generalised time that combines generalised times of both transport modes. Moreover, we analysed the effect of the systematic taste variation of households for accessibility, introducing interaction terms based on socio-demographic characteristics.

Usually, the comparison between models is based on statistical indicators like log of likelihood, likelihood ratio tests and rho-squared. In our work, we took a step forward and we analysed the capacity of the model to reproduce the observed market shares of location choices. When developing models for planning, it is important to reproduce correctly the observed market shares.

The main result of this study is that accessibility remains a significant variable in residential location choice models when we apply continuous accessibility. Their inclusion increases significantly the quality of the model and corrects the counterintuitive sign of some variables, like the housing price. In our application for Lyon, we observed that two variables had a counter-intuitive sign when we did not include accessibility or when we included the proximity to transport infrastructures defined as a binomial variable. All other estimations with accessibility measures give consistent results and parameters with stable signs. Moreover, the inclusion of systematic taste variation for accessibility using interaction terms proved to be essential. The modelling results were significantly better because households appreciate accessibility differently based on their socio-demographic characteristics.

With regard to the debate around the importance of accessibility in residential location choice, we support that transport and land-use modellers must include some sort of accessibility measure in their models, in line with Eliasson (2010) and Lee et al. (2010). The omission of accessibility could lead to erroneous results. When possible, and especially if the model has been developed for simulation purposes, the use of complex accessibility measures is preferable. In our analysis, the potential accessibility measure gave the best results. The analysis of the performance of the models to replicate the observations proved the superiority of the potential accessibility measure with a composite generalised time (this measurement method has been retained for the following chapters). This measure minimises systematic errors between observed and simulated choices. The inclusion of households' characteristics is useful from an analysis and simulation point of view. This is because the interaction terms capture some systematic taste variation of the households' preferences.

From a policy point of view, the results showed the importance of land use and individual components of accessibility. Decision-makers should take into consideration that residential policies focusing on the supply of activities around residential areas could be more efficient at attracting new households than the construction of new transport infrastructures. While it is important to have fast mobility solutions, households search for opportunities near their location of residence. In our case study, more than half of the moved households chose a location in central areas, where various activities and amenities are easily accessible. This is why central areas are attractive for households, posing a risk of eviction of sensitive groups. There are social groups that appreciate good accessibility, but they do not always have the financial means to pay the premium to live in highly accessible areas.

Do new and relocating firms have different preferences for accessibility?

While empirical works addressing the impact of accessibility on location choices are numerous, few papers have studied the differences in the valuation of accessibility between new and relocating firms. New and relocating firms do not place the same importance on location attributes (Duranton and Puga, 2001; Holl, 2004a; Manjón-Antolín and Arauzo-Carod, 2011). Accessibility can be one of these attributes, but empirical literature has not analysed its effect thoroughly.

Both new and relocating firms search for profit maximising locations (Barrios et al., 2006). Nevertheless, we can distinguish three major differences between them; (i) the local information at their disposal, (ii) the life cycle stage and (iii) the "ecosystem". When a firm is relocating within a geographical area, it holds key local information about the economic environment and market. Consequently, it can make an informed evaluation of the different location possibilities (Manjón-Antolín and Arauzo-Carod, 2011) and take the best possible decision for its economic activity. On the contrary, a new business, which enters a new territory, misses this local information. This will drive the firm to not take many risks related to location. Regarding the life stage of a firm, the creation and the relocation of an economic establishment are two distinct events in the life course of a firm (Holl, 2004a). A migrating firm searches for a better location than an actual one (Nguyen et al., 2013), while a new firm searches for the best location to house its first premises. Lastly, when a firm is already in a geographical area, it has an "ecosystem" (Moore, 1993), meaning the established network with workers, suppliers and clients from a spatial perspective. This dependency on the pre-existing network poses a restriction; relocating firms do not migrate far from the previous location because this acts as a pull factor (Bodenmann and Axhausen, 2012; Van Dijk and Pellenbarg, 2000).

Using this framework to analyse the potential differences in terms of preference for accessibility, we can form the following assumption. A new firm should be more risk averse and try to locate in an area where it can minimise all potential risks. Therefore, new firms are expected to localise themselves in central areas where accessibility is high and risks are low. A relocating firm can take more risks (Holl, 2004a), but has more constraints because of its "ecosystem" and its life stage.

Given this theoretical approach, **chapter II** provided empirical evidence on potential differences between new and relocating firms for eight different sectors. The objective was to highlight the differences of preferences for accessibility between the new and relocating economic establishments[12] belonging to the same economic sector. In that respect, we relied on a location choice model at a micro-level for eight economic sectors (Manufacturing, Construction, Wholesale, Retail, FIRE[13], Front Office Business Services, Back Office Services and Health), where we distinguished between new and relocating firms. The analysed creations and relocations occurred during 2005-2011. As explanatory factors, the model included measures for localisation and urbanisation effects, social environment, premises' price, existence of an economic activity zone and accessibility. Accessibility was included using three measures: (i) potential/gravity accessibility to population with a composite generalised time, (ii) proximity to transport infrastructure and (iii) preference for centrality. The comparison was based on elasticities to identify important differences.

The results showed that the location choices of new and relocating establishments differ significantly. The hypothesis that creations are more sensitive to accessibility to population and relocations more sensitive to transport infrastructure generally holds with some exceptions. A classification of the sectors can be made in three groups based on the differences of preferences for accessibility (table II.10). In the first group, we have the sectors that have a routine and production-oriented economic activities (Manufacturing, Construction, Wholesale and Back Office business services). Creations of these sectors search for locations with good accessibility,

[12] We use the terms economic establishment and firm interchangeably

[13] Finance, Insurance and Real Estate

in order to minimise all potential risks. However, migrating establishments select areas close to transport infrastructure, notably motorways. They do not value accessibility because it is not important for their economic activity. These locations offer good market access (Holl, 2004a). In the second group, we have the high-order services (FIRE and Front Office business services) that require daily face-to-face interaction and information exchange (Shearmur and Alvergne, 2002) and accessibility is important for their activity. This is why migrating establishments, which have the previous experience of the area, choose areas that have better accessibility in general (better accessibility to population/better proximity to transport infrastructure). In the third group, we have sectors that have a specific economic activity that depends highly on client interaction (Retail and Health). Creations of these sectors have higher preference for accessibility to population, in line with our initial hypotheses. Lastly, migrations in general are highly sensitive to the distance to the previous location, independently of the economic sector, confirming the findings of the literature (Elgar et al., 2015; Holl, 2004a; Manjón-Antolín and Arauzo-Carod, 2011). Firms want to be in proximity to their established "ecosystem".

For model developers, these results mean that, depending on the economic sector, the use of a different model for creations and migrations should be considered, or at least integrate this differentiation with interaction terms. For policy-making, these results imply that a transport investment, even if it increases the attractiveness of a location, it can decrease the relative attractiveness of nearby locations, causing business migrations in a certain radius. The applied policies should pay attention to this fact and give the appropriate incentives so that an investment leads to a general economic growth.

What is the impact of accessibility on the location choices of the business services?

All economic establishments do not give the same importance to accessibility. The previous chapter demonstrated this difference based on firm events (creation, relocation). However, it is not the only characteristic that can distinguish this preference. Establishments having different functions can also have different preferences (Ota and Fujita, 1993).

An economic sector that presents such a heterogeneity is the business services sector. In this sector, we can identify businesses that can be considered to be knowledge intensive, such as Research and Development or Business Administration. These activities have a high degree of final demand orientation and a high need for proximity, in order to exchange information on a daily basis. This type of firms can be considered as Front Office. At the same time, in the business services sector, we can also find establishments that have more routine function. This type of establishments can be decentralised, because they do not need spatial proximity and face-to-face interaction. Such activities can be call centres or rental agencies. The decentralisation of such Back Office functions is called “back officing” of routine functions (Ota and Fujita, 1993).

Given the different location choice logic between those two functions of the business services sector, **chapter III** aimed to analyse in more detail the differences of the preferences for accessibility. The objective was to compare the location choices between those two functions using statistical measures, as in chapter II. However, in chapter III the analysis was more thorough because the comparison was between only two sectors. The comparison was based on the results of a location choice model using statistical tests such as t-test, to compare the statistical difference of the obtained model parameters and elasticities to compare the volume of these differences. The analysis studied creations and relocations of Front Office and Back Office services that occurred during 2005-2011. The model included measures for localisation and urbanisation effects, social environment, premises’ price, existence of an economic activity zone and accessibility using the same four indicators used in chapter II.

The results demonstrate that the functional characteristics of the economic establishments are reflected on their preferences for accessibility. The location choices of Front Office services, the economic activity of which relies on daily interaction and information exchange, follow ideally the traditional trade-off between premises’ price and accessibility. Comparing to Back Office services, Front Office services have a stronger preference for central areas, where they can enjoy very good accessibility to population and where location externalities are strong. On the contrary, Back Office establishments can offer their services at a distance because direct proximity is less important for their economic activity. Therefore, they prefer peripheral areas in Economic Activity Zones, where they have easy access to motorways.

The results of this chapter can help the decision- and policy-makers to construct better policies, especially in transport. Investments on transport infrastructures or definitions of Economic Activity Zones can attract different types of economic establishments. An activity zone close to a dense and populated area with high accessibility to population would be more appropriate for a different type of services than an Economic Activity Zone in the outskirts. Lastly, we support that future works should take into account the distinction that this article proposes based on the function of the business services sector because differences can be significant.

Renters vs owners. How has the preference for accessibility evolved in residential location choices?

Previous chapters have demonstrated the importance of accessibility for location choices of households and firms. However, as accessibility rises more and more owing to economic growth and new transport infrastructure, people can enjoy more and more activities at the same or a lower generalised cost. This increase of accessibility might decrease its importance for both households and firms. Here we focus on residential location choices. During the second half of the 20th century, consistently with the hypothesis of the decreasing preference for accessibility (Alonso, 1964), the boost of accessibility due to the democratisation of the automobile led to the suburbanisation, the urban sprawl and longer trips to work. Various works confirmed this tendency, with observations from the USA (Baum-Snow, 2007) and Europe (Axhausen, 2008; Cabrera Delgado and Bonnel, 2016; Garcia-Lopez, 2012; Levinson, 2008). While the suburbanisation process is still underway, literature has identified a revival of city centres during the last 20 years (Buzar et al., 2007; Deka, 2018; Melia et al., 2018; Moos, 2013; Rérat, 2015). Behind this revival we can find the newer generations (millennials), which prefer to live in central locations with good accessibility, where the public transport supply is satisfactory, at the expense of spacious but car-dependent suburbs (Deka, 2018; Melia et al., 2018). What seems to drive this residential choice is the preference for accessibility and the advantages that highly accessible locations offer (Thomas et al., 2015).

Behind those two diverging location choice preferences that emerge, a characteristic that qualifies households is the housing occupation status. Owners, on the one hand, are less sensitive to accessibility and choose to buy a residence at peripheral areas. On the other hand, renters, which are mostly young households, select to rent an apartment in locations that offer high accessibility, despite the constant increase of accessibility and the high price of these locations.

Despite these observations, there are not many empirical works that have analysed the temporal evolution of residential location choice preferences to confirm those tendencies. Given this context, in **chapter IV,** we posed the following questions: How different are the preferences for accessibility between the renters and the owners? More importantly, how do the preferences of those two groups evolve over time? In case of accessibility increase, do the two groups have the same reaction? If not, how do owners and renters adapt their location choices? To respond to these questions, we relied on a residential location choice model using the same group of variables as in chapter I. The observations derive from the disaggregated census data of 1999, 2008 and 2013. We calculated the respective elasticities for each analysis year, in order to analyse potential tendencies. The urban area of Lyon is an interesting case study because, during that period, there was a significant increase of employment and population, there were substantial investments in public transport supply and, on the contrary, there were no important investments in car infrastructure.

The temporal analysis of the location choice preferences of renters and owners proved to be relevant. First, the results concerning the preference for accessibility between the two types of housing occupation status confirm our initial hypothesis that renters are more sensitive to accessibility. Second, the temporal analysis demonstrated a more interesting aspect of the preferences. Despite the fact that renters already had a higher preference for accessibility during the first analysis period (1999), the difference between owners and renters increases over time. Owners show a decreasing preference for accessibility, while renters an increasing one. What can explain this difference is the profile of the renters, which seem to correspond to the millennial generation identified in the literature (Deka, 2018; Melia et al., 2018; Myers et al., 2019). They are more likely to be younger, without a stable job and without children. The temporal analysis also showed that renters own cars at a decreasing rate, which could explain their constant preference for zones with good public transport services.

Our results suggest that transport planning and land-use policies should be more integrated, to take into consideration those temporal dynamics. A future radical and general improvement of accessibility levels is expected to take place thanks to new mobility solutions like Mobility as a Service or automated vehicles (Meyer et al., 2017). As we demonstrated, the improvement of accessibility can increase the polarisation between suburban and central areas. The former ones would be even more diffused and the latter ones even denser. Planners should anticipate those changes and channel household choices through housing supply, in order to avoid undesirable situations like the social exclusion and the segregation of a certain type of households. The anticipation could be successful through the integration of temporal dynamics into the modelling procedure. Methods combining different cross-sectional data can capture these temporal variations of the preferences and help in estimating better predictive models (Rezaei and Patterson, 2016).

General contribution

Throughout this thesis dissertation, the framework of accessibility and the use of discrete choice models permitted the grouping of the location choices of households and firms under the same problematic. The general conclusions that we draw from this exercise can be related to practical issues, methodological guidelines and policy instructions.

At practical level, the various constraints (data, models) limited the analysis presented in this thesis to the application of potential accessibility as the most complex measure. However, this procedure highlighted the importance of aggregation between transport modes. While there are various aggregation methods proposed in the literature, they are not relevant for use in the context of location choice models. They are either difficult to be estimated or they are not consistent with the concept of accessibility. The use of composite generalised times is theoretically solid (consistent with the accessibility concept, see chapter I) and proved that the obtained results are superior. The analysis performed in chapter 1 provided the appropriate confidence to apply the same methodology in all the other chapters as well. This process proved that the analyst has to use all the available data, to find the most appropriate methodology in the objective to bypass practical limits.

Applying accessibility measures entails methodological choices that need to be justified. Regarding the individual dimension for both households and firms, the in-depth analysis of potential differences between different households and firms showed that there exist important differences. As for households, it has demonstrated that all types are more or less sensitive to accessibility, with students, workers with no stable employment status and renters being the most sensitive ones. With regard to firms, while accessibility has a positive influence, it depends on the characteristics of the firms (economic sector, function, creation/relocation) and on the type of accessibility. Capturing this kind of individual taste variations can lead to better modelling results. Their inclusion needs attention, in order to avoid problems of overfitting caused by endogeneity issues. Regarding the land-use component of accessibility, the choice between population and employment depends on the analysed agent. The choice affects the construction of the measure and the interpretation of the results.

At policy instruction level, this dissertation highlights the multidimensionality of accessibility. The historically established approach of accessibility only from a transport perspective has led the applied policies to focus on the supply of mobility solutions. Many projects were realized that helped the spread of fast, motorised and individual transport means. However, the negative effects of such policies like pollution, congestion, long commute distances are known and visible in many cities. Accessibility should be integrated into the discussion as the combined effect of transport and land use, taking into account the component of individuality. Accessibility changes can be driven by land-use changes, which might be slower, but their effects are more resilient and in line with the guidelines of sustainable development (trip optimisation, active modes of transport, etc.). However, the increase of accessibility might result in effects like gentrification. Such effects, depending on their strength, might penalise certain social groups. The applied policies should integrate countermeasures aiming to avoid such negative effects.

Limits and perspectives

A PhD thesis is a limited in time research project. This means that one cannot address all possible issues related to the subject. However, the shortcomings of this project can be a starting point for future research. Therefore, identifying them and giving future directions is important for every research project.

In this dissertation, the focus was placed on a single yet interesting case study, the urban area of Lyon. While the obtained results of the different chapters are in line with previous research works, confirming the transferability of the results in other spatial contexts is important for their generalisation. Each city has some particularities of its own and it is important to demonstrate them in empirical studies. Both accessibility and LUTI models have strong empirical foundations. Systematic confirmation of their consistency is important for their generalisation.

Another limit concerns the accessibility measurement. In this work, accessibility is approached in an as general as possible manner for comparison reasons between different models. However, in order to ameliorate location choice models, the development of more specific accessibility measures can be envisaged (e.g. accessibility to shopping, leisure or by employment type). Furthermore, more disaggregated accessibility measures can be studied, like a utility-based measure (e.g. logsum). In this study, the data was insufficient to calculate consistent utility-based accessibility measures at the spatial scale of location choice models. Nevertheless, if data is available, the benefit of complex accessibility measures like logsum should be evaluated in comparison to traditional measures like opportunity-based measures. Furthermore, the recent approach of accessibility as a measure of general life satisfaction (Chaloux et al., 2019; Lättman et al., 2016) can give new ideas of integration of dimensions, like the satisfaction of making a trip (Chaloux et al., 2019), or access to information and communication technologies. These dimensions affect the perceived accessibility and influence the quality of life (Wee et al., 2013).

The applied modelling method throughout the thesis could also be a point of criticism. The four chapters adapted a MultiNomial logit model to explain the location choice models of households and firms. Nevertheless, even if chapter IV applied a more suitable modelling method, there are theoretically more appropriate methods to model choices in a spatial context, like mixed spatially correlated models (Bhat and Guo, 2004). Such methods integrate better the correlation of the errors between alternatives. This integration of correlation seems the most promising. However, its application needs a special development of models that was outside the scope of this thesis. Another future modelling direction can focus on the analysis of simultaneous choice models for residential location, employment location, car ownership, etc. Those choices are interdependent and the direction of causality remains unclear in the literature. Future findings can give new insights on these complex choice issues.

Lastly, a drawback of this dissertation, limited to location choice models, was the absence of an analysis on the simulation results of a LUTI model. While these concerns are relevant, they were out of the thesis' scope mostly due to time constraints. They are, however, considered to be integrated in future works.

Concluding, we need to mention that, as the notion of accessibility is complex and in constant evolution, newer and more innovative definitions might surface in the future. The norms of living, working and socialising constantly change and influence how people interact and shape their environment (telecommuting, e-shopping, e-governance, online dating, streaming services, etc.). This means that the need for spatial interactions with employment, leisure and shopping opportunities might evolve. New sources of data can help us to understand these complex spatial relations that seem to emerge. However, in order to fully take advantage of these capabilities, we need a solid theoretical framework that incorporates accessibility as a more general concept. Future theoretical works should move in this direction.

Bibliography

Acheampong, R.A., Silva, E.A., 2015. Land use – transport interaction modeling: A review of the literature and future research directions. J. Transp. Land Use 8, 11–38.

Aissaoui, H., 2016. Choix de localisation résidentielle des ménages en milieu urbain. Les apports récents des modèles de choix discrets en présence d'un nombre élevé d'alternatives. Université Lumière Lyon 2. PhD dissertation.

Aissaoui, H., Bouzouina, L., Bonnel, P., 2015. Choix de localisation résidentielle, entre contraintes du marché et préférences individuelles : application à l'aire urbaine de Lyon 1999. Rev. d'Economie Régionale Urbaine 4, 629–656.

Alamá-Sabater, L., Artal-Tur, A., Navarro-Azorín, J.M., 2011. Industrial location, spatial discrete choice models and the need to account for neighbourhood effects. Ann. Reg. Sci. 47, 393–418. https://doi.org/10.1007/s00168-010-0383-7

Alonso, W., 1964. Location and land use : toward a general theory of land rent. Harvard University Press, Cambridge, Mass.

Anas, A., Arnott, R., Small, K.A., 1998. Urban Spatial Structure. J. Econ. Lit. 36, 1426–1464. https://doi.org/http://www.jstor.org/stable/2564805

Arauzo-Carod, J.M., Liviano-Solis, D., Manjon-Antolin, M., 2010. Empirical Studies in Industrial Location: an Assessment of Their Methods and Results. J. Reg. Sci. 50, 685–711. https://doi.org/10.1111/j.1467-9787.2009.00625.x

Axhausen, K.W., 2008. Accessibility: Long-term perspectives. J. Transp. Land Use 1, 5–22. https://doi.org/10.5198/jtlu.v1i2.66

Axhausen, K.W., Froelich, P., Tschopp, M., 2011. Changes in Swiss accessibility since 1850. Res. Transp. Econ. 31, 72–80. https://doi.org/10.1016/j.retrec.2010.11.010

Baptista, R., Mendonça, J., 2010. Proximity to knowledge sources and the location of knowledge-based start-ups. Ann. Reg. Sci. 45, 5–29. https://doi.org/10.1007/s00168-009-0289-4

Baraklianos, I., Bouzouina, L., Bonnel, P., 2018a. The impact of accessibility on the location choices of the business services . Evidence from Lyon urban area. Région et Développement 48, 85–104.

Baraklianos, I., Bouzouina, L., Bonnel, P., Aissaoui, H., 2018b. Does the accessibility measure influence the results of residential location choice modelling? Transportation. https://doi.org/s11116-018-9964-6

Barrios, S., Görg, H., Strobl, E., 2006. Multinationals' location choice, agglomeration economies, and public incentives. Int. Reg. Sci. Rev. 29, 81–107. https://doi.org/10.1177/0160017605281516

Baum-Snow, N., 2007. Did Highways Cause Suburbanization? Q. J. Econ. 122, 775–805. https://doi.org/10.1162/qjec.122.2.775

Baumont, C., 2009. Spatial effects of urban public policies on housing values. Pap. Reg. Sci. 88, 301–326. https://doi.org/10.1111/j.1435-5957.2008.00200.x

Beaudry, C., Schiffauerova, A., 2009. Who's right, Marshall or Jacobs? The localization versus urbanization debate. Res. Policy 38, 318–337. https://doi.org/10.1016/j.respol.2008.11.010

Ben-Akiva, M.E., Bierlaire, M., Mcfadden, D., Walker, J., 2015. Discrete Choice Analysis. Preliminary draft version.

Ben-akiva, M.E., Bowman, J.L., 1998. Integration of an activity-based model system and a residential location model. Urban Stud. 35, 1131–1153.

Ben-Akiva, M.E., Lerman, S., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand, MIT Press Series in Transportation Studies. MIT Press, Cambridge, USA.

Bhat, C.R., Carini, J., Misra, R., 1999. Modeling the Generation and Organization of Household Activity Stops. Transp. Res. Rec. J. Transp. Res. Board 1676, 153–161. https://doi.org/10.3141/1676-19

Bhat, C.R., Guo, J.Y., 2004. A mixed spatially correlated logit model: Formulation and application to residential choice modeling. Transp. Res. Part B Methodol. 38, 147–168. https://doi.org/10.1016/S0191-2615(03)00005-5

Bhat, C.R., Handy, S., Kockelman, K., Mahmassani, H., Chen, Q., Weston, L., Handy, S., Mahmassani, H., Weston, L., 2000. Urban Accessibility Index: Literature Review. Texas.

Bhat, C.R., Handy, S., Kockelman, K., Mahmassani, H., Weston, L., Gopal, A., Srour, I., 2002. Development of an urban accessibility index: Formulations, aggregation and application. Texas.

Blijie, B., 2005. The impact of accessibility on residential choice : Empirical results of a discrete choice model, in: 45TH Congress of the European Regional Science Association.

Bodenmann, B.R., 2011. Location choice of firms with special emphasis on spatial accessibility. ETH Zürich. PhD dissertation.

Bodenmann, B.R., Axhausen, K.W., 2012. Destination choice for relocating firms: A discrete choice model for the St. Gallen region, Switzerland. Pap. Reg. Sci. 91, 319–341. https://doi.org/10.1111/j.1435-5957.2011.00389.x

Boisjoly, G., El-Geneidy, A.M., 2016. Daily fluctuations in transit and job availability: A comparative assessment of time-sensitive accessibility measures. J. Transp. Geogr. 52, 73–81. https://doi.org/10.1016/j.jtrangeo.2016.03.004

Boiteux-Orain, C., Huriot, J.-M., 2002. Modéliser la suburbanisation. Rev. d'Économie Régionale Urbaine février, 73. https://doi.org/10.3917/reru.021.0073

Bonnel, P., 2004. Prévoir la demande de transport. Presses de l'Ecole Nationale des Ponts et Chaussées, Paris.

Bonnel, P., Bouzouina, L., Monchambert, G., 2013. Croissance urbaine et accessibilité : intérêts et limites de l'approche par les densités dans une perspective de modélisation de l'usage du sol. Cybergeo Eur. J. Geogr.

Bouzouina, L., 2008. Ségrégation spatiale et dynamiques métropolitaines. Université Lumière Lyon 2. PhD dissertation.

Bouzouina, L., Aissaoui, H., Baraklianos, I., 2018. Dynamiques des choix de localisation des ménages. Focus sur l'aire urbaine de Lyon et ses quartiers Zus entre 1999 et 2013, in: Mobilité Résidentielle Des Habitants Des Quartiers Prioritaires. Rapport 2017. Observatoire national de la politique de la ville, pp. 108–124.

Bouzouina, L., Cabrera-Delgado, J., Emmerich, G., 2014. Inégalités d'accessibilité à l'emploi en Transport Collectif Urbain : Deux décennies d'évolutions en banlieue lyonnaise. Rev. d'Economie Régionale Urbaine 1, 33–61.

Bouzouina, L., Nicolas, J.-P., 2015. Du « tout voiture » au « Vélo'v », cinquante ans d'évolution du système de transport lyonnais, in: Rosales-Montano, S., Aguilera-Belanger, A., Mignot, D., Terral, L. (Eds.), L'expérience Métropolitaine Lyonnaise. FNAU, Paris, pp. 43–44.

Brueckner, J.K., Thisse, J.F., Zenou, Y., 1999. Why is central Paris rich and downtown Detroit poor? An amenity-based theory. Eur. Econ. Rev. 43, 91–107. https://doi.org/10.1016/S0014-2921(98)00019-1

Buczkowska, S., de Lapparent, M., 2014. Location choices of newly created establishments: Spatial patterns at the aggregate level. Reg. Sci. Urban Econ. 48, 68–81. https://doi.org/10.1016/j.regsciurbeco.2014.05.001

Bunel, M., Tovar, E., 2014. Key Issues in Local Job Accessibility Measurement: Different Models Mean Different Results. Urban Stud. 51, 1322–1338. https://doi.org/10.1177/0042098013495573

Buzar, S., Ogden, P., Hall, R., Haase, A., Kabisch, S., 2007. Splintering Urban Populations : Emergent Landscapes of Reurbanisation in Four European Cities. Urban Stud. 44, 651–677.

Cabrera-delgado, J., 2013. Quelle prise en compte des dynamiques urbaines dans la prévision de la demande de transport ? Université Lumière Lyon 2. PhD dissertation.

Cabrera Delgado, J., Bonnel, P., 2016. Level of aggregation of zoning and temporal transferability of the gravity distribution model: The case of Lyon. J. Transp. Geogr. 51, 17–26. https://doi.org/10.1016/j.jtrangeo.2015.10.016

Callon, 2005. La cote annuelle des valeurs vénales immobilières et foncières, Editions C. ed. Surgères, France.

Cao, X., Mokhtarian, P.L., Handy, S.L., 2009. Examining the impacts of residential self-selection on travel behaviour: A focus on empirical findings. Transp. Rev. 29, 359–395. https://doi.org/10.1080/01441640802539195

Carlton, D.W., 1983. The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables. Rev. Econ. Stat. 65, 440–449.

Carpenter, J., Verhage, R., 2014. Lyon city profile. Cities 38, 57–68. https://doi.org/10.1016/j.cities.2013.12.003

Cascetta, E., Cartenì, A., Montanino, M., 2016. A behavioral model of accessibility based on the number of available opportunities. J. Transp. Geogr. 51, 45–58. https://doi.org/10.1016/j.jtrangeo.2015.11.002

Cerema, 2014. Zones d'activité économique en périphérie : les leviers pour la requalification. Lyon. Available at https://www.adcf.org/files/co18005314-zae-etude-cerema.pdf

Chaloux, N., Boisjoly, G., Grisé, E., El-Geneidy, A., & Levinson, D., 2019. I only get some satisfaction: Introducing satisfaction into measures of accessibility. Paper presented at the 98th Annual Meeting of the Transportation Research Board, Washington D.C., USA.

Chen, J., Chen, C., Timmermans, H., 2008. Accessibility Trade-Offs in Household Residential Location Decisions. Transp. Res. Rec. J. Transp. Res. Board 2077, 71–79. https://doi.org/10.3141/2077-10

Cheshire, P.C., 2006. Resurgent Cities, Urban Myths and Policy Hubris: What We Need to Know. Urban Stud. 43, 1231–1246. https://doi.org/10.1080/00420980600775600

Cho, E.J., Rodriguez, D. a, Song, Y., 2008. The role of employment subcenters in residential location decisions. J. Transp. Land Use 2, 121–151. https://doi.org/10.5198/jtlu.v3i1.194

Christaller, W., 1933. Die centralen Orte in Suddeutschland. Gustav Fischer, Jena. English edition: Christaller W (1967) Central places in Southern Germany (trans: Baskin CW). Englewood Cliffs, New Jersey.

Coppola, P., Nuzzolo, A., 2011. Changing accessibility, dwelling price and the spatial distribution of socio-economic activities. Res. Transp. Econ. 31, 63–71. https://doi.org/10.1016/j.retrec.2010.11.009

Cordera, R., Coppola, P., dell'Olio, L., Ibeas, Á., 2017. Is accessibility relevant in trip generation? Modelling the interaction between trip generation and accessibility taking into account spatial effects. Transportation. 44, 1577–1603. https://doi.org/10.1007/s11116-016-9715-5

Dalvi, M.Q., Martin, K.M., 1976. The measurement of accessibility: Some preliminary results. Transportation. 5, 17–42.

de Bok, M., 2007. Infrastructure and Firm Dynamics: A micro-simulation approach. TU Delft. PhD dissertation

de Bok, M., Van Oort, F., 2011. Agglomeration economies, accessibility and the spatial choice behavior of relocating firms. J. Transp. Land Use 4, 5–24. https://doi.org/10.5198/jtlu.v4i1.144

De Palma, A., Motamedi, K., Picard, N., Waddell, P., 2007. Accessibility and Environmental Quality : Inequality in the Paris Housing Market. Eur. Transp. / Trasp. Eur. 36, 47–74.

De Palma, A., Motamedi, K., Picard, N., Waddell, P., 2005. A model of residential location choice with endogenous housing prices and traffic for the Paris region. Eur. Transp. / Trasp. Eur. 31, 67–82.

Deboosere, R., Kerrigan, D.J., Wachsmuth, D., El-Geneidy, A.M., 2019. Location, location and professionalization: a multilevel hedonic analysis of Airbnb listing prices and revenue. Reg. Stud. Reg. Sci. 6, 143–156. https://doi.org/10.1080/21681376.2019.1592699

Deka, D., 2018. Are millennials moving to more urbanized and transit-oriented counties ? J. Transp. Land Use 11, 443–461. https://doi.org/10.5198/jtlu.2018.1345

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 A Policy Pract. 40, 163–180. https://doi.org/10.1016/j.tra.2005.05.002

Dubé, J., Brunelle, C., Legros, D., 2016. Location Theories and Business Location Decision : A Micro-Spatial Investigation of a Non-Metropolitan Area in Canada. Rev. Reg. Stud. 46, 143–170.

Duranton, G., Puga, D., 2005. From sectoral to functional urban specialisation. J. Urban Econ. 57, 343–370. https://doi.org/10.1016/j.jue.2004.12.002

Duranton, G., Puga, D., 2001. Nursery Cities: Urban diversity, process innovation, and the life-cycle of products. Am. Econ. Rev. 91, 1545–1477. https://doi.org/10.1257/aer.91.5.1454

Duranton, G., Turner, M.A., 2012. Urban growth and transportation. Rev. Econ. Stud. 79, 1407–1440. https://doi.org/10.1093/restud/rds010

El-Geneidy, A.M., Cerda, A., Fischler, R., Luka, N., 2011. The use of accessibility measures to evaluate the impacts of transportation plans: An application in Montréal, Québec. Can. J. Urban Res. Can. Plan. Policy 20, 81–104.

Elgar, I., Farooq, B., Miller, E.J., 2015. Simulations of firm location decisions: Replicating office location choices in the Greater Toronto Area. J. Choice Model. 17, 39–51. https://doi.org/10.1016/j.jocm.2015.12.003

Elgar, I., Farooq, B., Miller, E.J., 2009. Modeling Location Decisions of Office Firms. Transp. Res. Rec. J. Transp. Res. Board 2133, 56–63. https://doi.org/10.3141/2133-06

Eliasson, J., 2010. The Influence of Accessibility on Residential Location, in: Pagliara, F., Preston, J., Simmons, D. (Eds.), Residential Location Choice: Models and Applications. Springer, pp. 137–164. https://doi.org/10.1007/978-3-642-12788-5

Eliasson, J., Mattsson, L.-G., 2000. A model for integrated analysis of household location and travel choices. Transp. Res. Part A Policy Pract. 34, 375–394. https://doi.org/10.1016/S0965-8564(99)00038-5

Ellison, B.G., Glaeser, E.L., Kerr, W.R., 2010. What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. Am. Econ. Rev. 100, 1195–1213. https://doi.org/10.1257/aer.100.3.1195

Florida, R., 2004. The rise of the creative class and how it's transforming work, leisure, community and everyday life, Paperback. ed.

Fox, J., Daly, A., Hess, S., Miller, E., 2014. Temporal transferability of models of mode-destination choice for the Greater Toronto and Hamilton Area. J. Transp. Land Use 7, 41. https://doi.org/10.5198/jtlu.v7i2.701

Frenkel, A., Shefer, D., Koschatzky, K., Walter, G.H., 2001. Firm Characteristics, Location and Regional Innovation: A Comparison Between Israeli and German Industrial Firms. Reg. Stud. J. Reg. Stud. Assoc. 35, 415–429. https://doi.org/10.1080/00343400120058424

García-Albertos, P., Picornell, M., Salas-Olmedo, M.H., Gutiérrez, J., 2018. Exploring the potential of mobile phone records and online route planners for dynamic accessibility analysis. Transp. Res. Part A Policy Pract. https://doi.org/10.1016/j.tra.2018.02.008

Garcia-Lopez, M.A., 2012. Urban spatial structure, suburbanization and transportation in Barcelona. J. Urban Econ. 72, 176–190. https://doi.org/10.1016/j.jue.2012.05.003

Geurs, K.T., Ritsema van Eck, J.R., 2001. Accessibility measures: review and applications. Evaluation of accessibility impacts of land-use transport scenarios, and related social and economic impacts. RIVM report, 408505006. National Institute of Public Health and the Environment, Bilthoven

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

Glaeser, E.L., Kahn M. E., 2004. Sprawl and Urban Growth, in: Henderson, J. V., Thisse, J.E. (Eds.), Handbook of Regional and Urban Economics, Cities and Geography. Elsevier, Amsterdam, pp. 2481–2527.

Glaeser, E.L., Kallal, H.D., Scheinkman, J.A., 1992. Growth in cities. J. Polit. Econ. 100, 1126–1152.

Glaeser, E.L., Kolko, J., Saiz, A., 2001. Consumer city. J. Econ. Geogr. 1, 27–50. https://doi.org/10.1093/jeg/1.1.27

Guevara, C., Ben-Akiva, M.E., 2006. Endogeneity in Residential Location Choice Models. Transp. Res. Rec. J. Transp. Res. Board 60–66. https://doi.org/10.3141/1977-10

Guevara, C.A., 2015. Critical assessment of five methods to correct for endogeneity in discrete-choice models. Transp. Res. Part A Policy Pract. 82, 240–254. https://doi.org/10.1016/j.tra.2015.10.005

Guevara, C.A., Ben-Akiva, M.E., 2013. Sampling of alternatives in Multivariate Extreme Value (MEV) models. Transp. Res. Part B Methodol. 48, 31–52. https://doi.org/10.1016/j.trb.2012.11.001

Guo, J.Y., 2004. Addressing spatial complexities in residential location choice models. University of Texas.

Guo, J.Y., Bhat, C.R., 2007. Operationalizing the concept of neighborhood: Application to residential location choice analysis. J. Transp. Geogr. 15, 31–45. https://doi.org/10.1016/j.jtrangeo.2005.11.001

Hägerstrand, T., 1970. What About People in Regional Science ? 9th Eur. Congr. Reg. Sci. Assoc. 24, 7–21.

Halden, D., 2002. Using accessibility measures to integrate land use and transport policy in Edinburgh and the Lothians. Transp. Policy 9, 313–324. https://doi.org/10.1016/S0967-070X(02)00017-3

Hamersma, M., Heinen, E., Tillema, T., Arts, J., 2015. Residential moving intentions at highway locations : The trade-off between nuisances and accessibility in the Netherlands. Transp. Res. Part D 35, 130–141. https://doi.org/10.1016/j.trd.2014.11.017

Handy, S.L., Niemeier, D.A., 1997. Measuring accessibility: An exploration of issues and alternatives. Environ. Plan. A 29, 1175–1194. https://doi.org/10.1068/a291175

Hansen, W.G., 1959. How Accessibility Shapes Land Use. J. Am. Inst. Plann. 25, 73–76. https://doi.org/10.1080/01944365908978307

Haque, M.B., Choudhury, C.F., Hess, S., n.d. Understanding differences in residential preferences between ownership and renting with consideration of changes over time: a case study of London, Working Paper.

Hayter, R., 1997. The dynamics of industrial location. The factory, the firm and the production system. John Wiley and Sons, Chichester, UK. https://doi.org/10.1002/9781118786352.wbieg0444

Holl, A., 2012. Market potential and firm-level productivity in Spain. J. Econ. Geogr. 12, 1191–1215. https://doi.org/10.1093/jeg/lbr030

Holl, A., 2004a. Start-ups and relocations: Manufacturing plant location in Portugal. Pap. Reg. Sci. 83, 649–668. https://doi.org/10.1007/s10110-004-0218-y

Holl, A., 2004b. Manufacturing location and impacts of road transport infrastructure: Empirical evidence from Spain. Reg. Sci. Urban Econ. 34, 341–363. https://doi.org/10.1016/S0166-0462(03)00059-0

Homocianu, M., 2009. Modélisation de l'intéraction transport-urbanisme - Choix residentiels des ménages dans l'aire urbaine de Lyon. Université Lumière Lyon 2. PhD dissertation.

Hu, L., 2017. Job accessibility and employment outcomes: which income groups benefit the most? Transportation. 44, 1421–1443. https://doi.org/10.1007/s11116-016-9708-4

Hu, L., Wang, L., 2017. Housing location choices of the poor: does access to jobs matter? Hous. Stud. 3037, 1–25. https://doi.org/10.1080/02673037.2017.1364354

Ibeas, Á., Cordera, R., Dell'Olio, L., Coppola, P., 2013. Modelling the spatial interactions between workplace and residential location. Transp. Res. Part A Policy Pract. 49, 110–122. https://doi.org/10.1016/j.tra.2013.01.008

Inoa, I.A., Picard, N., de Palma, A., 2015. Effect of an Accessibility Measure in a Model for Choice of Residential Location, Workplace, and Type of Employment. Math. Popul. Stud. 22, 4–36. https://doi.org/10.1080/08898480.2014.925318

INSEE, 2013. Recensement de la population.

Jacobs, J., 1969. The economy of the cities, Vintage. ed. New York.

Kasraian, D., Maat, K., Stead, D., van Wee, B., 2016. Long-term impacts of transport infrastructure networks on land-use change: an international review of empirical studies. Transp. Rev. 36, 772–792. https://doi.org/10.1080/01441647.2016.1168887

Klein, N.J., Smart, M.J., 2017. Millennials and car ownership: Less money, fewer cars. Transp. Policy 53, 20–29. https://doi.org/10.1016/j.tranpol.2016.08.010

Koenig, G., 1980. Indicators of urban accessibility: Theory and application. Transportation. 9, 145–172. https://doi.org/10.1007/BF00167128

Koenig, G., 1974. Théorie économique de l'accessibilité urbaine. Rev. économique 25, 275–297. https://doi.org/10.3406/reco.1974.408144

Kryvobokov, M., Bouzouina, L., 2014. Willingness to pay for accessibility under the conditions of residential segregation. Int. J. Strateg. Prop. Manag. 18, 101–115. https://doi.org/10.3846/1648715X.2013.864342

Kwan, M.-P., 1998. Space-Time and Integral Measures of Individual Accessibility: A Comparative Analysis Using a Point-based Framework. Geogr. Anal. 30, 191–216. https://doi.org/10.1111/j.1538-4632.1998.tb00396.x

Lättman, K., Friman, M., Olsson, L.E., 2016. Perceived Accessibility of Public Transport as a Potential Indicator of Social Inclusion. Soc. Incl. 4. https://doi.org/10.17645/si.v4i3.481

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

Levinson, D., 2008. Density and dispersion: The co-development of land use and rail in London. J. Econ. Geogr. 8, 55–77. https://doi.org/10.1093/jeg/lbm038

Lowry, I.S., 1964. A model of metropolis. California.

Manjón-Antolín, M.C., Arauzo-Carod, J.M., 2011. Locations and relocations: Determinants, modelling, and interrelations. Ann. Reg. Sci. 47, 131–146. https://doi.org/10.1007/s00168-009-0360-1

Maroto, A., Zofío, J.L., 2016. Accessibility gains and road transport infrastructure in Spain: A productivity approach based on the Malmquist index. J. Transp. Geogr. 52, 143–152. https://doi.org/10.1016/j.jtrangeo.2016.03.008

Marshall, A., 1890. Principles of Economics. Online Libr. Lib. 1–627. https://doi.org/10.1057/9781137375261

Martín-Barroso, D., Núñez-Serrano, J.A., Velázquez, F.J., 2017. Firm heterogeneity and the accessibility of manufacturing firms to labour markets. J. Transp. Geogr. 60, 243–256. https://doi.org/10.1016/j.jtrangeo.2017.03.008

McFadden, D., 1977. Modelling the choice of residential location. Spat. Interact. Theory Plan. Model.

McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior, in: Zarembka, P. (Ed.), Frontiers in Econometrics. Academic Press, New York, pp. 105–142. https://doi.org/10.1108/eb028592

Mejia-Dorantes, L., Paez, A., Vassallo, J.M., 2012. Transportation infrastructure impacts on firm location: The effect of a new metro line in the suburbs of Madrid. J. Transp. Geogr. 22, 236–250. https://doi.org/10.1016/j.jtrangeo.2011.09.006

Melia, S., Chatterjee, K., Stokes, G., 2018. Is the urbanisation of young adults reducing their driving ? Transp. Res. Part A 118, 444–456. https://doi.org/10.1016/j.tra.2018.09.021

Melo, P.C., Graham, D.J., Levinson, D.M., Aarabi, S., 2016. Agglomeration, accessibility and productivity: Evidence for large metropolitan areas in the US. Urban Stud. 0042098015624850. https://doi.org/10.1177/0042098015624850

Meyer, J., Becker, H., Bösch, P.M., Axhausen, K.W., 2017. Autonomous vehicles: The next jump in accessibilities? Res. Transp. Econ. 62, 80–91. https://doi.org/10.1016/j.retrec.2017.03.005

Milakis, D., Kroesen, M., van Wee, B., 2018. Implications of automated vehicles for accessibility and location choices: Evidence from an expert-based experiment. J. Transp. Geogr. 68, 142–148. https://doi.org/10.1016/j.jtrangeo.2018.03.010

Miller, H.J., 1999. Measuring Space-Time Accessibility Benefits within Transportation Networks: Basic Theory and Computational Procedures. Geogr. Anal. 31, 1–26.

Mills, E.S., 1967. Transportation and patterns of urban development: An Aggregative Model of Resource Allocation in a Metropolitan Area. Am. Econ. Rev. 57, 197–210. https://doi.org/10,1007/978-1-4684-4001-0

Millsap, A., 2016. Location choice in early adulthood: Millennials versus Baby Boomers. Pap. Reg. Sci. 97, S139–S167. https://doi.org/10.1111/pirs.12240

Moore, J.F., 1993. Predators and prey: A new ecology of competition. Harv. Bus. Rev. 71, 75–86. https://doi.org/10.1111/gec3.12359

Moos, M., 2013. “Generationed” space: Societal restructuring and young adults’ changing residential location patterns. Can. Geogr. 58, 11–33.

Mulley, C., 2014. Accessibility and Residential Land Value Uplift: Identifying Spatial Variations in the Accessibility Impacts of a Bus Transitway. Urban Stud. 51, 1707–1724. https://doi.org/10.1177/0042098013499082

Muth, R.F., 1969. Cities and housing: the spatial pattern of urban land use. University of Chicago Press, Chicago, IL.

Myers, D., Lee, H., Simmons, P.A., 2019. Cohort Insights into Recovery of Millennial Homeownership after the Great Recession. J. Hous. Econ. https://doi.org/10.1016/j.jhe.2019.01.004

Nguyen, C.Y., Sano, K., Tran, T.V., Doan, T.T., 2013. Firm relocation patterns incorporating spatial interactions. Ann. Reg. Sci. 50, 685–703. https://doi.org/10.1007/s00168-012-0523-3

Nicolas, J.-P., 2010. SIMBAD : un outil pour intégrer le développement durable dans les politiques publiques, in: Jean-Pierre, A. (Ed.), Modéliser La Ville : Formes Urbaines et Politiques de Transport. Economica, Paris, pp. 120–164.

Niedzielski, M.A., Eric Boschmann, E., 2014. Travel Time and Distance as Relative Accessibility in the Journey to Work. Ann. Assoc. Am. Geogr. 104, 1156–1182. https://doi.org/10.1080/00045608.2014.958398

Oakil, A.T.M., Manting, D., Nijland, H., 2016. Determinants of car ownership among young households in the Netherlands: The role of urbanisation and demographic and economic characteristics. J. Transp. Geogr. 51, 229–235. https://doi.org/10.1016/j.jtrangeo.2016.01.010

Ota, M., Fujita, M., 1993. Communication technologies and spatial organization of multi-unit firms in metropolitan areas. Reg. Sci. Urban Econ. 23, 695–729.

Padeiro, M., 2013. Transport infrastructures and employment growth in the Paris metropolitan margins. J. Transp. Geogr. 31, 44–53. https://doi.org/10.1016/j.jtrangeo.2013.05.007

Pellenbarg, P.H., van Wissen, L., Van Dijk, J., 2002. Firm relocation: state of the art and research prospects, SOM research school. Groningen.

Plaut, P.O., 2006. The intra-household choices regarding commuting and housing. Transp. Res. Part A Policy Pract. 40, 561–571. https://doi.org/10.1016/j.tra.2005.10.001

Portnov, B.A., Axhausen, K.W., Tschopp, M., Schwartz, M., 2011. Diminishing effects of location ? Some evidence from Swiss municipalities , 1950 – 2000. J. Transp. Geogr. 19, 1368–1378. https://doi.org/10.1016/j.jtrangeo.2011.07.017

Priemus, H., Nijkamp, P., Banister, D., 2001. Mobility and spatial dynamics: An uneasy relationship. J. Transp. Geogr. 9, 167–171. https://doi.org/10.1016/S0966-6923(01)00007-2

Rérat, P., 2015. Le retour des villes. Espac. Popul. sociétés 3.

Rezaei, A., Patterson, Z., 2016. Preference stability in household location choice: Using cross-sectional data from three censuses. Res. Transp. Econ. 67, 44–53. https://doi.org/10.1016/j.retrec.2016.11.006

Rosales-Montano, S., Aguilera-Belanger, A., Mignot, D., Terral, L., 2015. L'expérience métropolitaine lyonnaise. FNAU, Paris.

Rosenthal, S.S., Strange, W.C., 2003. Geography, Industrial Organization, and Agglomeration. Rev. Econ. Stat. 85, 377–393. https://doi.org/10.1162/003465303765299882

Schirmer, P.M., Van Eggermond, M. a. B., Axhausen, K.W., 2014. The role of location in residential location choice models: a review of literature. J. Transp. Land Use 7, 3–21. https://doi.org/10.5198/jtlu.v7i2.740

Sener, I.N., Pendyala, R.M., Bhat, C.R., 2011. Accommodating spatial correlation across choice alternatives in discrete choice models: an application to modeling residential location choice behavior. J. Transp. Geogr. 19, 294–303. https://doi.org/10.1016/j.jtrangeo.2010.03.013

Shearmur, R., Alvergne, C., 2002. Intrametropolitan Patterns of High-order Business Service Location: A Comparative Study of Seventeen Sectors in Ile-de-France. Urban Stud. 39, 1143–1163.

Srour, I., Kockelman, K., Dunn, T., 2002. Accessibility Indices: Connection to Residential Land Prices and Location Choices. Transp. Res. Rec. 1805, 25–34. https://doi.org/10.3141/1805-04

Sweet, M.N., 2014. Do firms flee traffic congestion? J. Transp. Geogr. 35, 40–49. https://doi.org/10.1016/j.jtrangeo.2014.01.005

Sytral, 2007. Enquête ménages déplacements 2006 de l'aire métropolitaine lyonnaise : Pricipaux résultats. Lyon.

Tenkanen, H., 2017. Capturing time in space: Dynamic analysis of accessibility and mobility to support spatial planning with open data and tools. Unigrafia, Helsinki.

Thomas, E., Serwicka, I., Swinney, P., 2015. Urban demographics Why people live where they do.

Van Dijk, J., Pellenbarg, P.H., 2000. Firm relocation decisions in The Netherlands: An ordered logit approach. Pap. Reg. Sci. 79, 191–219. https://doi.org/10.1007/s101100050043

Vincent-Geslin, S., Pochet, P., Ortar, N., Bonnel, P., Bouzouina, L., 2017. Je t'aime… moi non plus. EspacesTemps.net.

Von Thünen, J.H., 1842. Der isolirte Staat in Beziehung auf Landwirthschaft und Nationalökonomie, 2nd ed. Rostock.

Waddell, P., 2010. Modeling Residential Location in UrbanSim, in: Pagliara, F., Preston, J., Simmons, D. (Eds.), Residential Location Choice: Models and Applications. Springer, pp. 165–180. https://doi.org/10.1007/978-3-642-12788-5

Waddell, P., 1998. An Urban Simulation Model for Integrated Policy Analysis and Planning: Residential Location and Housing Market Components of UrbanSim, in: 8th World Conference on Transport Research. Antwerp, Belgium.

Waddell, P., 1996. Accessibility and Residential Location: The Interaction of Workplace, Residential Mobility, Tenure, and Location Choices. Lincoln L. Inst. TRED Conf. 1–19.

Wang, Y., Chen, B.Y., Yuan, H., Wang, D., Lam, W.H.K., Li, Q., 2018. Measuring temporal variation of location-based accessibility using space-time utility perspective. J. Transp. Geogr. 73, 13–24. https://doi.org/10.1016/j.jtrangeo.2018.10.002

Washington, S.P., Karlaftis, M.G., Mannering, F.L., 2011. Statistical and econometric methods for transportation data analysis, 2nd editio. ed, Journal of Chemical Information and Modeling. CRC Press, New York. https://doi.org/10.1017/CBO9781107415324.004

Weber, A., 1909. Über den Standort der Industrie, English translation: Theory of the Location of Industries, The Univer. ed.

Wee, B. Van, 2015. Toward a new generation of land use transport interaction models : a viewpoint. J. Transp. Land Use 3, 1–10.

Wee, B. Van, Geurs, K.T., Chorus, C., 2013. Information, communication, travel behavior and accessibility. J. Transp. Land Use 6, 1–16. https://doi.org/10.5198/jtlu.v6i3.282

Wegener, M., 2004. Overview of land-use transport models, in: Handbook of Transport Geography and Spatial Systems, Vol 5. Pergamon/Elsevier Science, pp. 127–146. https://doi.org/10.1007/s10654-011-9614-1

Wegener, M., Fürst, F., 1999. Land-Use Transport Interaction: State of the Art. Dortmund.

Weisbrod, G.E., Lerman, S.R., Ben-akiva, M.E., 1980. Tradeoffs in residential location decisions: transportation versus other factors. Transp. Policy Decis. Mak. 1, 13–26.

Wilson, A.G., 1998. Land Use transport interaction models: past and future. J. Transp. Econ. Policy 32, 3–26.

Wilson, A.G., 1971. Family of spatial interaction models, and associated developments. Environ. Plan. 3, 1–32. https://doi.org/10.1068/a030001

Wilson, A.G., 1967. A statistical theory of spatial distribution models. Transp. Res. 1, 253–269. https://doi.org/https://doi.org/10.1016/0041-1647(67)90035-4

Wilson, R.R., Krizek, K.J., El-Geneidy, A.M., 2007. How Stable are Preferences for Neighborhood Characteristics & Accessibility? Analyzing Residential Relocation Decisions, in: 11th World Conference on Transport Research Society. Berkley CA, p. 27.

Zahavi, Y., 1974. Travel time budgets and mobility in urban areas. Report FHW PL-8183, Washington, D: US Department of Transportation.

Zondag, B., de Bok, M., Geurs, K.T., Molenwijk, E., 2015. Accessibility modeling and evaluation: The TIGRIS XL land-use and transport interaction model for the Netherlands. Comput. Environ. Urban Syst. 49, 115–125. https://doi.org/10.1016/j.compenvurbsys.2014.06.001

Zondag, B., Pieters, M., 2005. Influence of Accessibility on Residential Location Choice. Transp. Res. Rec. J. Transp. Res. Board, Advances in Spatial Science 1902, 63–70. https://doi.org/10.3141/1902-08

Abstract

This PhD thesis has as objective to examine the place and the importance of accessibility in location choice models of households and firms, two key construction elements of Land-Use Transport Interaction models. More specifically, the aim is to analyse the effect of various methodological choices from a theoretical and empirical point of view in order to give some answers to theoretical, methodological, empirical and policy issues. Having as a case study the urban area of Lyon and using discrete models to explain the location choices of households and firms, four research papers comprise the main work of this dissertation.

In the first paper, the objective was to analyse the effect of the accessibility measure on the results of residential location choice model. While accessibility has always been important at theoretical level, at empirical level, some works questioned its importance, considering other location characteristics as more influential. This paper examines whether different accessibility measurement methods can lead to divergent results. The conclusion is that accessibility is an indispensable variable for residential location choice models and the conclusion remains the same whatever is the measure. Without accessibility, the model gives inconsistent results. Complex accessibility measures give better results, especially for predictions, but simple measures are also relevant for residential location choices modelling. The choice highly depends on the objectives of the application especially if the model is to be used for simulation.

In the second paper, the objective was to analyse the differences of accessibility preferences between new and relocating firms. Accessibility is one of the most important attributes of a location choice of an economic establishment. However, even if it seems intuitive, works analysing any differences between creations and relocations are scarce. Using data from eight economic sectors and comparing creations to relocations, the results demonstrate that the effect of accessibility differs between in the same economic sector. This difference depends on the type of economic activity of the sector but also on the type of accessibility. With some exceptions, creations are more sensible to accessibility to population while relocations are more sensible to proximity to transport infrastructure. Transport policies aiming to attract economic activities could provoke relocations from nearby locations decreasing the economic activity inside the same wider area.

In the third paper, the objective was to evaluate the difference of the accessibility impact on the location choices of firms of the business services sector. Distinguishing between Front Office and Back Office business services in a location choice model, the results demonstrate that the effect of accessibility differs between economic subsectors. Consistent with the type of their economic activity, Front Office services prefer highly accessible location with good transport infrastructures where location externalities are strong. Face-to-face interactions are important for their activity. Back Office services are sensitive only to the proximity to motorways.

In the fourth paper, the objective was to analyse the temporal evolution of the preferences for accessibility for residential choices. More and more people choose to buy a residence at the suburbs taking advantage of the accessibility increase. At the same time, young households, the so called millennials, choose to rent in central areas. Distinguishing between renters and owners, the analysis of the elasticities for 1999, 2006 and 2013 confirm the initial intuition. Renters were always more sensitive to accessibility to employment. More importantly, renters show an increasing preference for accessibility during the analysis period, while the owners the opposite. We suggest that planners and model developers should integrate temporal dynamics into their models in order to anticipate better future tendencies.

The main result of the PhD thesis is that accessibility has an undeniable central role in location choice models and in LUTI in general and that applying accessibility measures entails methodological choices that need to be justified. Analyses showed that accessibility has an important role for location choices of both households and firms, but for different reasons. For households, it has demonstrated that all types of households are more or less sensitive to accessibility, with students, workers with no stable employment status and renters being the most sensible. This means that locations with good accessibility attract households to move in and they are ready to pay a premium in order to live into these locations. For firms, while accessibility has positive influence, it depends on the characteristics of the firms (economic sector, function, creation/relocation) and on the type of accessibility.

Résumé

La présente thèse de doctorat a pour objectif d'examiner la place et l'importance de l'accessibilité dans les modèles de choix de localisation des ménages et des entreprises. Ces modèles sont deux éléments clés de la conception et la construction des modèles d'interaction transport – usages du sol. Il s'agit, plus précisément, d'analyser l'effet de divers choix méthodologiques d'un point de vue théorique et empirique afin de donner des réponses à des interrogations tout aussi théoriques, méthodologiques, empiriques et politiques. Pour y répondre, quatre articles de recherche constituent le travail principal de la thèse. Les quatre travaux produits ont pour objet l'étude de l'aire urbaine lyonnaise et utilisant des modèles de choix discrets pour expliquer les choix de localisation.

Dans le premier article, l'objectif est d'analyser l'effet de la mesure d'accessibilité sur les résultats du modèle de choix de localisation résidentielle. Alors que l'accessibilité a toujours été importante au niveau théorique, certains travaux relativisent son importance au niveau empirique, considérant que d'autres attribues de localisation sont plus influentes. Cet article analyse si différentes méthodes de mesure de l'accessibilité peuvent conduire à des résultats divergents.

La conclusion principale est que l'accessibilité est une variable indispensable pour les modèles de choix localisation résidentielle et ce quelle que soit la mesure. Sans la variable de l'accessibilité, le modèle donne des résultats incohérents. De plus, les mesures complexes donnent de meilleurs résultats, en particulier pour les prévisions, mais des mesures simples sont également pertinentes pour la modélisation des choix résidentiels. Le choix, entre les deux types de mesure, dépend fortement des objectifs de l'application, en particulier si le modèle doit être utilisé pour la simulation.

Dans le deuxième article, l'objectif est d'analyser les différences de préférences en matière d'accessibilité entre les entreprises nouvellement créées et les entreprises qui se relocalisent. L'accessibilité est l'un des facteurs les plus importants du choix de localisation d'une entreprise. Cependant, même si cela semble intuitif, les travaux analysant les différences entre les créations et les relocalisations sont peu nombreux. En utilisant des données pour huit secteurs d'activités économiques et en confrontant les créations aux relocalisations, les résultats démontrent que l'effet de l'accessibilité diffère d'un secteur économique à l'autre. Cette différence dépend du type d'activité économique du secteur mais aussi du type d'accessibilité. Avec quelques exceptions, les créations sont plus sensibles à l'accessibilité à la population alors que les relocalisations montrent une sensibilité pour la proximité aux infrastructures de transport. Les politiques de transport visant à attirer les activités économiques pourraient avoir pour effet contraire des relocalisations depuis des sites proches ; incitant davantage le déplacement des entreprises plutôt que des créations. Par conséquent, l'activité économique, dans la zone considérée, ne profitera pas pleinement de tous les biens faits de ces politiques.

Dans le troisième article, l'objectif est d'évaluer la différence d'impact de l'accessibilité sur les choix de localisation des entreprises du secteur des services aux entreprises. En distinguant les services entre Front Office et Back Office selon leur fonction dans un modèle de choix de localisation, les résultats démontrent que l'effet de l'accessibilité diffère d'un sous-secteur économique à l'autre. En fonction de leur type d'activité économique, les services de Front Office préfèrent une localisation facilement accessible dotée de bonnes infrastructures de transport où les effets de localisation sont importants. Les services de back office, quant à eux, ne sont sensibles qu'à la proximité des autoroutes.

Dans le quatrième papier, l'objectif est d'analyser l'évolution temporelle des préférences en matière d'accessibilité des choix résidentiels. De plus en plus de ménages choisissent d'acheter une résidence en banlieue, profitant ainsi de l'augmentation de l'accessibilité. En opposition, les jeunes ménages, appelés aussi «*millennials*», choisissent de louer dans les zones centrales. Distinguant les locataires des propriétaires, l'analyse des élasticités de périodes 1999, 2006 et 2013 confirme l'intuition initiale qui est que les locataires sont plus sensibles à l'accessibilité à l'emploi. Plus important encore, la préférence des locataires évolue et croit au cours de la période analysée, alors que celle des propriétaires évolue de façon inverse. Nous suggérons aux planificateurs et aux concepteurs de modèles d'intégrer la dynamique temporelle dans leurs modèles afin d'anticiper au mieux les tendances futures.

Les principales conclusions de la thèse sont que l'accessibilité joue un rôle indéniable dans les modèles de choix de localisation et dans les LUTI en général et que l'application de mesures d'accessibilité implique des choix méthodologiques qui doivent être justifiés. Les analyses ont montré également que l'accessibilité joue un rôle important dans les choix de localisation des ménages et des entreprises, mais pour des raisons différentes. Pour les ménages, il a été démontré que tous les types de ménages sont plus ou moins sensibles à l'accessibilité, les étudiants, les travailleurs sans statut d'emploi stable et les locataires étant les plus sensibles. Cela signifie que les endroits facilement accessibles incitent les ménages à s'y installer. De plus, ces derniers sont prêts à payer un supplément pour pouvoir y vivre.

Pour les entreprises, si l'accessibilité a une influence positive, elle varie selon leurs caractéristiques tels que le secteur économique, la fonction, s'il s'agit d'une création ou d'une relocalisation et/ou du type d'accessibilité.