

LOCATIONAL CHOICE AND PREFERENCES WITH COSTLY RELOCATION

A STUDY ON HOUSEHOLD RESIDENTIAL SORTING IN THE
EINDHOVEN REGION

Master Thesis

Marco Gatto, 2517750

August 18th, 2013

Supervisor: Prof.dr. Jan Rouwendal

Second Supervisor: dr. Mark van Duijn

Vrije Universiteit Amsterdam

MSc Spatial, Transport and Environmental Economics

Abstract

The spatial distribution of households has long been considered by urban economists, as people are expected to reveal their preferences for local amenities when choosing their residential location. The thesis analyses household residential sorting in the Eindhoven area through a horizontal sorting model. We extend the analytical framework of Bayer et al. (2004) by relaxing the assumption of perfect mobility of agents. Moving to a new dwelling entails a mixture of financial and psychological costs that households prefer avoiding. This study demonstrates that households, especially the higher educated and the older ones, are extremely reluctant to move away from their original location. On average, we find that a household would be willing to pay about 60% more for a standard house if that allows him to avoid moving to a different municipality. The paper provides also estimates of the MWTP for locational characteristics for different types of households. Our analysis confirms that preferences for urban amenities vary substantially with the socio-demographic characteristic of households.

Table of Contents

1. Introduction	2
2. Context	3
3. Urban and Regional Economic Theory: an Overview	5
3.1. Cities and production	6
3.2. Cities and consumption	7
3.3. Amenities, rents and growth	8
4. The Locational Choice of Households	10
4.1. The Equilibrium sorting model	12
4.2. Econometric Specification of the Residential Sorting Model	14
5. Data and Descriptives	19
6. Estimation Results	25
6.1. Results from the first-stage regression of the residential sorting model	25
6.2. Results from the second-stage regression of the residential sorting model	30
6.3. Marginal Willingness to Pay	31
7. Conclusions	34
References	37

1. Introduction

In an era of diminishing transportation costs and increasing accessibility to goods and services, households' well-being has still a lot to do with factors related to the space and the living environment. Cities within the same region, as well as neighbourhoods within the same city, differ among each other under several aspects, so that many are the factors that potentially contribute to the attractiveness of a location as a residential place for households. Transport needs and cost of living are surely of primary importance for dwellers, as emphasized by seminal urban economic models (Muth, 1969, Glaeser and Kahn, 2004). But also the availability of good schools, open space and in general public goods has always been taken into high considerations both by households and urban experts (Tiebout, 1956, Bayer et al. 2007, Klaiber et al., 2007). More recently, however, some authors have argued that households, especially the wealthier ones, are increasingly attracted also by other types of amenities. For instance, Brueckner et al. (1999) infer that affluent residents have strong preferences not only for short commutes and spacious houses, but also for amenities such as architectural beauty, social atmosphere, leisure amenities, etc. Similarly, Glaeser et al. (2001) postulates that consumption amenities such as bars, restaurants, shops, recreational activities, etc. are normal goods, and that skilled households will be more and more concentrated in cities that can offer a vast choice of such amenities.

In this study we seek to shed light on the multifaceted and intricate topic of households' locational choice. More precisely, we analyse households' residential sorting behaviour within a localized housing/labour market, i.e. the South-east North Brabant region. Our study focuses on two main questions pertaining to households' sorting behaviour. First, we want to see how and to what extent certain amenities influence the locational choice of different categories of households. Second, we want to test empirically whether residents can actually move freely across locations, or whether instead moving costs, i.e. the cost of relocating to a new dwelling, affect their decision.

There is typically a high degree of heterogeneity in the way households value locational characteristics, which is one of the reasons why it often results very difficult to explain or predict the distribution of households across the urban space. Therefore, we employ a model of locational choice that can effectively deal with the heterogeneity that arises across locations, and especially across households. Specifically, we use an equilibrium sorting model with additive random utility (Bayer et al., 2004, van Duijn and Rouwendal, 2013) to see how the willingness to pay for locational attributes vary with respect to households' sociodemographic characteristics. We extend the standard horizontal sorting models' framework by relaxing the assumption of perfect mobility of agents. Moving costs can be intended as the combination of psychological and financial costs that households incur when relocating to a new dwelling. The idea is that households might be restrained from choosing their optimal location if moving is costly. As frictions in the housing market generate distortions in the observed sorting equilibrium, one can expect that overlooking moving costs would bias the estimation of households' demand for locational characteristics.

The locational characteristics considered in this paper are housing price, natural amenities (natural and water coverage), urban amenities (number of bars, restaurant, shops, etc) and various measures

of accessibility. Household characteristics we control for are instead educational level, marital status, presence of children and age of the family's head. We are interested in investigating the preference profile of one group of residents in particular, i.e. of highly-educated households. There exists a vast body of literature providing evidences of the ultimate role that human capital play nowadays in thriving regional development (Acemoglu and Angrist, 2001, Florida, 2004, Moretti, 2012). Such issues are extremely relevant also for the region we consider in our analysis. The South-east North Brabant economy is indeed highly specialized in knowledge intensive productions such as high-tech manufacturing and HTSM. This implies that future regional growth potentials largely depend on the ability of local stakeholders to retain and attract human capital intensive firms, and the skilled labour force needed.

In the light of that, our analysis is potentially relevant both for scholars and policy makers. For the former, as it links to the literature on consumers' preferences for urban amenities, embedding the interesting and important extension concerning moving costs. For the former, because it provides a number of meaningful welfare measures that city planners can beneficially use to design sound and effective spatial plans and investments.

The study is structured as follows. Section 2 describes the main features characterizing SNB regional economy. Section 3 provides some insights on relevant theories and contributions developed in the field of urban and regional economics. Section 4 is devoted to illustrate the econometric framework used in the analysis, and contains a detailed discussion on the equilibrium sorting model we refer to. Section 5 provides some elucidations on the data used, together with descriptive statistics on our sample. Estimation results are reported and discussed in section 6. Section 7 concludes the paper.

2. Context

The South-east North Brabant (SNB), or simply Eindhoven region, is a geographical area located in the southern part of the Netherlands, bordered by the Limburg province in the east and the Belgium's Antwerp province in the south. The region's capital, Eindhoven, with a population of around 218,000 inhabitants, is the largest city in the North Brabant province, and the fifth largest throughout the country. At the heart of a rich and densely populated area of Europe, the region lays in a favourable and well-connected geographical location, crossroads of prominent transport networks and industrial clusters such as those of the Dutch Randstad, the Flemish urban diamond, and the German Ruhr. Moreover, due to its strong industrial tradition and its top technology sector, South-East Brabant contributes significantly to the Dutch economy, vaunting the highest national shares of exports, private R&D expenditure and patents produced (Brainport Eindhoven, 2005). Despite lacking the charming historical and cultural infrastructures available in many Dutch cities (Amsterdam, Den Haag, Utrecht, etc.), SNB remains a very attractive region for skilled workers thanks to his unique and highly competitive business environment, see Figure 1. To better understand the SNB-specific factors of success, this section provides an overview of the regional economy, with a focus on the industrial strengths and the specialization pattern characterizing the local business environment.

South-East Brabant as a Knowledge Hotspot

SNB represents one of the main national cluster of knowledge-intensive activities, and is by many recognized as the industrial leader among all Holland's Southern regions. SNB's strong industrial

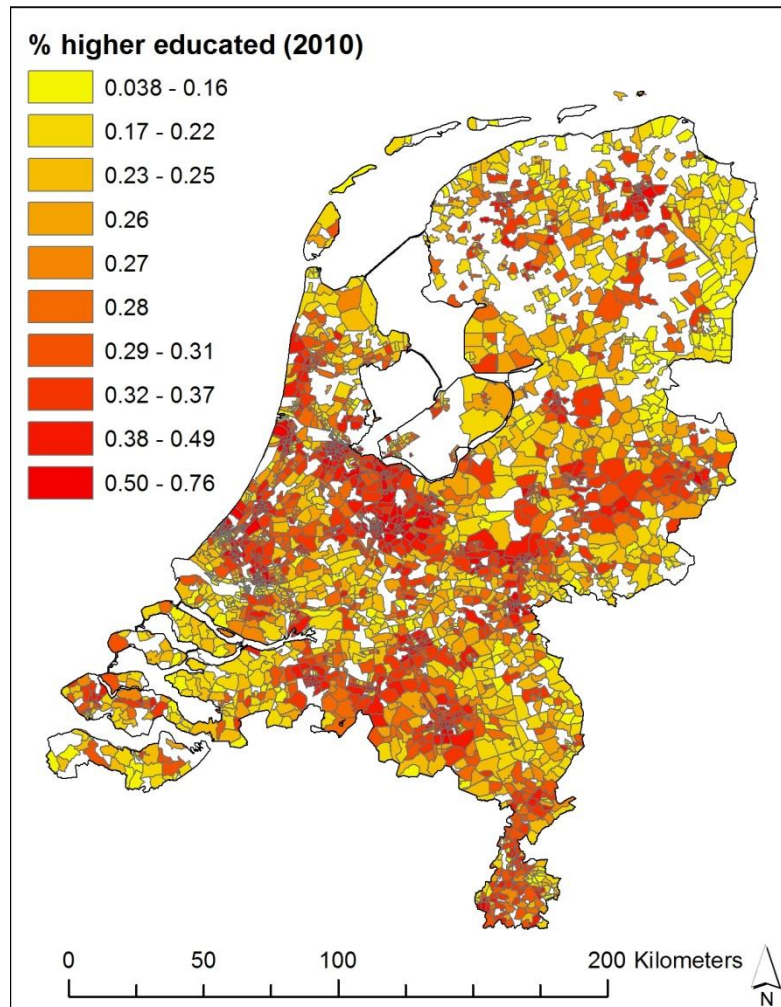


Figure 1. Percentage of highly-educated individuals in Dutch municipalities. Source: CBS (2010)

position is built upon a system of various types of public and private knowledge-based entities. Local employment and investments are driven in the first place by big international companies (Philips, ASML, DAF Trucks, etc.) which played a pivotal role for the regional industrial development during both its earlier stages and the more recent post-industrial reorganization. These are then supported by a number of knowledge-intensive small and medium sized companies (SMEs) which, beyond functioning as multinationals' satellites activities, maintain the local industrial system highly innovative and dynamic. On the top of that, numerous are the R&D institutions that concur to provide the business environment with innovative solutions and technologies, and to bring new talents and ideas into the region.¹ Among others, Automotive, LifeTec & Health, Food, High Tech Systems & Materials (HTSM) and Design are the sectors that mostly contribute to the regional economy. These, which count for a good 16.1% of all regional jobs, can be seen as the

¹ R&D institutions vary from Public Educational Institutes (DAE, TU/e, FONTYS), to Private R&D centres (e.g. HTC Eindhoven, DPI), and semi-private or independent knowledge institutions (TNO, ESI, TTAI, ECN, FOM).

“key cornerstones” of the local economy, and are recognized by many as the most strategic sectors for SNB's future employment and productivity growth (Brainport Eindhoven, 2012).

Given their common knowledge-intensive profile, this type of productions has spawned a very dynamic and fertile business environment, with strong cross-industrial synergies and collaborations. HTSM enterprises such as Philips, ASML and NXP, for instance, while serving the world's market with their top technology manufactures, are closely connected with many SMEs and R&D institutions that, through their abilities and highly specialized activities, play a fundamental role along the value-chain of knowledge-intensive productions: the former, by supplying high-tech intermediate products and solutions; the latter, by collecting, selecting and interpreting information and knowledge that are necessary for the development of new products and processes (the DAE's industrial design academy and the TU/e and FONTYS polytechnics being good examples of that). All this conveys to clustered, vertically integrated networks of local actors that continuously exchange information and knowledge, carrying forward a process of interactive learning extremely capable of producing technological and processing innovations, and thereby new value to be sold on the global market.

A number of indicators confirm SNB's strength in innovation and technology. The region contributes by some 30% to overall national industrial employment and by 40% to national value added in manufacturing, with around 70,000 jobs in industry (whose more than half is employed in high- and medium-tech industries). A good 9.3% of SNB's Gross Regional Product is spent on R&D (whose 7.9% is private), a share that represents 30% of all spending on R&D made annually in Netherlands (45% if only private spending is counted). Moreover, SNB has more registered ICT patents than any other region in Europe (43% of all Dutch patents), and was named by the Intelligent Community Forum, a renowned think tank on social and economic development, the 2011 world's most intelligent community.²

Although the competition coming from the world's emerging economies is increasingly undermining the industrial position of most Western countries, SNB has still good possibilities to flourish in the future if investments on its existing strengths will be opportunely made. The system of ties and skills that are embedded in the SEB's industrial cluster represents a unique throughout the world, and the ultimate asset for any knowledge economy. It is thus essential for SNB to leverage all those factors that are decisive for regional growth and development. As the next section will try to highlight, such factors are related not only to the private sector's competitiveness, but also to all those amenities that affect the quality of living of a region.

3. Urban and Regional Economic Theory: An Overview

A vast body of literature has been produced in the field of modern urban and regional economics in the attempt of explaining spatial agglomerations and urban growth. As the wealth of regions and

² Data on R&D spending refer to the year 2007 (Brainport Development, 2012), while data on patents, regional jobs, industrial size, etc. were provided by Brainport (2005) and refer to the period 2003-2005 (a precise year of reference is not indicated in the report, see p.47).

nations is sharply dependent on their ability to attract people and production, understanding the drivers of factors' mobility has always been a crucial question in economics. Firms normally locate their production in the place that is most profitable to them, while people presumably choose the residence offering the best bundle of locational characteristics. Although different in their nature, these behaviours are inherently related to each other. As firms want to be close to consumers and workers, employment availability and conditions are of primary interest for dwellers. Urban amenities are important determinants of regional attractiveness, and this paper will provide interesting findings on such topic. At the same time, it is also interesting to have some insights on how the local labour market can play a major role for determining regional development and attractiveness. The aim of this section is thus to describe how production- and consumption-related factors interact over time and space, and how they concur together to determine regional growth and development.

3.1 Cities and production

Following the pivotal theorization made by Marshall (1920), the ability of cities to facilitate or stimulate productive activities is justified by the fact that agglomerations and density are essentially related to the gains generated by transport cost reductions. In principle, thus, the rise of geographic clusters can be thought as deriving from the willingness of economic agents to reduce the costs of transporting goods, people and ideas (Ellison et al., 2010). Hence, natural advantages, input-output sharing, thick labour markets and human capital externalities are some of the factors typically indicated by scholars as prompting spatial agglomeration of economic agents (Glaeser et al., 1995, Ellison et al., 2010).

Whatever their nature, agglomeration economies are considered by many as a powerful engine of growth for cities, and a concept that is fundamental for explaining the concentration of people and firms over space. In the last twenty years, however, the attention scholars putted on idea-based agglomeration economies have been remarkable. Particularly emphasized has been the influence that location-specific ideas and human capital have on local economies, and the mechanisms through which they can boost productivity and wealth of cities (see e.g. Glaeser and Gottlieb, 2009). Various empirical studies have shown indeed how localized scale externalities and information spillovers result being crucial especially for knowledge-based activities (see e.g. Henderson, 2003, and Duranton and Puga, 2001). This is because technology intensive industries succeed mainly according to their ability to develop new product and processes, or in one word “innovate”. Therefore, it is extremely important for them to operate within an environment that facilitate economic interactions, and thereby the spread of innovations and ideas.

Going one step further, if one reckons that the transmission of knowledge enhances productivity locally, it is then straightforward to assume that such effects increase with the local level of human capital. In other words, the more skilled the labour force, the higher the “quality” of the knowledge transferred among firms and workers, the more capable an urban system will be to innovate and grow. Of course the relationship between human capital and factor productivity may not sound as a novelty to many economists, given the extensive work that endogenous growth theory scholars have

provided on this topic. What urban economists have tried to describe is however the spatial dimension of such relationship, i.e. to what extent the level of human capital can influence factors' productivity through localized externalities. On the base of the idea that such external effects take place within a limited spatial area, many authors have shown that the level of human capital is indeed responsible of much of the cross-city variation in economic growth within a given country, that being measured indirectly by population-employment trends (Glaeser et al., 1995, Simon and Nardinelli, 2002), or directly by wage differentials among metropolitan areas (Rauch, 1991, Acemoglu and Angrist, 2001, Moretti, 2004).

More recently, an exhaustive study on the link between human capital, industrial specialization and regional development has been done by Moretti (2012). In his book “The New Geography of Jobs”, he describes the key role human capital has had within the process of industrial reorganization which the US economy went through in the last three decades. Moretti reckons how employment and wealth in the US is increasingly concentrated around what he calls “innovation hubs” (San Francisco, Seattle, Boston, etc.), i.e. those local economies that are specialized in human capital intensive industries such as high-tech, HTSM, life science, etc. Thanks to a relatively more skilled labour force, such cities turned out being extremely more attractive for workers and businesses of the innovation sectors. Following the movement of worker and firms, spatial clustering soon translated into rising productivity and wages, a trend that contrasted strongly with the decline experienced in the same period by low-skilled (former) manufacturing poles such as Detroit, Flint, Cleveland, etc. A particularly striking part of the story pertains to the effects that such knowledge based specialization pattern is producing in terms of wage differentials. Comparing high- with low-skilled cities, indeed, Moretti finds that the level of human capital enhances productivity and wages not only of the highly educated workers, but also of those with a lower level of education. In sum, his research emphasizes not the role of human capital in explaining cross-city differences in growth patterns, but also the fact that human capital spillovers affect the local environment on a very broad base.

3.2. Cities and consumption

As described in the previous section, human capital and agglomeration economies can lead, through localized productivity gains, to a substantial increase in metropolitan population and employment. However, it is hard to imagine that earning conditions are the only thing that matters when people choose the place where to live. Thinking about urban quality of life, for instance, who would rationally move to an unpleasant location, everything else being equal? Explaining migration trends and households' locational choice is of course much harder than it could appear. Not surprisingly, indeed, many economic models were developed upon the idea that factors other than production-related ones are in fact important determinants of locations' attractiveness.

Many authors have emphasized e.g. the importance of historical amenities such as monuments, fine architecture, and historical buildings. Brueckner et al. (1999), for instance, argue that cities equipped with such amenities, e.g. those with a pleasant historical centres, result being much more attractive to people, especially those with higher income. Van Duijn and Rouwendal (2013) come to

similar conclusions, demonstrating that in the Netherlands cultural heritage influences severely housing prices. Carlino and Saiz (2008) show how “beautiful cities”, i.e. those with more historical and other recreational amenities, are perceived as more attractive not only by visitors, but also by local dwellers. Exploiting the relationship between amenities and number of visitors, they also provide clear evidences on the fact that the availability of leisure-oriented amenities consistently predicts urban demographic growth. Thereby, they indicate such amenities as powerful policy instruments for attracting high-skilled households.

Natural amenities can also influence substantially the perceived attractiveness of cities. Some authors have shown e.g. that factors like winter temperature (Rappaport, 2009) and proximity to coast (Rappaport and Sachs, 2003) are among the most powerful predictors of population growth in the US. Also parks, open space, nice landscape, etc. are increasingly valued by consumers (Florida et al., 2011, Klaiber et al., 2007), and represent today an important policy instrument available to local governments.

Glaeser et al. (2001), moreover, stress the fact that the rise in average income experienced by most Western countries is expected to increase the demand for a wide range of consumption amenities, namely those classifiable as normal and luxury “goods”. Therefore, local public services such as education, security and fast transport infrastructures must be taken into serious consideration when thinking about regional development. The authors also claim that the rich variety of services and consumer goods available in large urban areas also represents a key determinant of cities’ attractiveness, as they spur not only the mere consumption of goods but also social contacts and human interactions. In their opinion, this holds especially for rich workers, as these are “less fettered by constraint on employment location”.

From this point of view, the comparative advantages enjoyed by larger conurbation in the provision of certain consumption amenities appear particularly relevant. Restaurants, theatres, concert venues, sport facilities, public infrastructures, specialized retailers etc., they all have substantial economies of scale. This implies that such amenities will be locally available only when the urban population, i.e. the number of potential customers, reaches some critical level. If one also assumes that wealthier, highly-educated people are in general more attracted by consumption amenities, then the mechanisms underlying urban growth will arguably be self-reinforcing. For instance, if a city can attract highly educated households because of some exogenous advantage, the local economy will (in principle) start expanding according to the prevailing agglomeration economies. Public facilities and consumption amenities will then follow such trend, making that location even more attractive for affluent households. Eventually, this type of spiral-effects suggests that many complementary forces can contribute to determine the economic development of a city, and that the endogenous forces related to human capital affects cities in terms of both consumption and production amenities.

3.3. Amenities, rents and growth

We have seen in the previous sections, a number of scale effects contribute to foster the

attractiveness of large urban areas. At the same time, however, also negative externalities such as crime, pollution, longer commutes, etc. are associated to large cities. Since the stock of infrastructures available locally cannot immediately accommodate the rise in urban population, various forms of congestion normally occur when migrating flows are sizable. A typical example of that pertains to non-traded goods such as housing, which are available only on a limited amount at any given point in time. With a relatively fixed supply, thus, only a small share of the immigrant population will find a vacant dwelling. The result will be that the housing market will react more through prices than through quantity increments.

The link between housing rents and city attractiveness is well rooted in urban economics, and is at the base of many works dealing with local amenities and urban quality of life. Roback (1982), for example, developed a neoclassical growth model that considers the interaction between wages, population and housing rents to construct a measure of urban quality of life. The model proposed by Roback relies on a fundamental condition of spatial economics, i.e. the existence of “spatial equilibrium”. The idea is that as firms and workers are perfectly mobile, utility is equalized across space, and individuals are indifferent across locations. In principle thus high nominal wages would be offset by negative urban attributes such as expensive housing or a low level of amenities. Among other things, such theoretical framework is useful for studying the relative importance that consumption or productive amenities have in determining city attractiveness.

Using data on growth in population, wages and house values at the city and metropolitan level, many authors have exploited the Roback's formulation to investigate the factors underlying the success experienced by some cities in the US during the last decades. Glaeser and Gottlieb (2006) e.g. observe that the real wage premium to live in big cities has been consistently declining over time³, and their interpretation is that such decline evidences an overall improvement in urban quality of life. At the same time they do not exclude that people might have started appreciating urban amenities more than in the past because of higher incomes and education level. Glaeser and Saiz (2003) extend the Roback's model by accounting for the local level of human capital. They find that, at the metropolitan level, human capital generates growth almost uniquely through productivity gains. However, they demonstrate that the link between population growth, human capital and urban amenities become much stronger if one considers cities instead of metropolitan areas. Contrasting findings for US metropolitan areas are provided by Shapiro (2005), who demonstrates that in the period 1940-1990, the share of college educated acted on employment growth only for about 60% through productivity growth, the rest being caused by growth in urban quality of life.

In sum, both consumption and productive amenities seem to be important factors of development for cities. However, economists have been quite inconclusive in specifying what type of

³ The wage premium can be defined as the additional remuneration required by workers to accept a job in a large city. This premium exists because of various disamenities (e.g. crime, congestion, etc.) that generally make larger cities relatively less attractive than smaller cities. When expressed in real terms, the nominal wage differential across locations is weighted for the corresponding local cost of living, namely house price.

agglomeration economies or consumption amenities are actually at stake. Despite that, human capital appears decisive under many aspects, especially in the way it prompts a series of endogenous mechanisms that simultaneously help local economies to thrive. Although our study has not much to do with issues related to regional growth or development, we believe that an overview of the mainline theories of regional and urban economics is useful for various reasons. Firstly, and most importantly, because it documents how relevant consumption, recreational and other types of local amenities are considered to be increasingly important for urban residents. Secondly, since it helps the reader acknowledge the important role human capital plays nowadays for the economic performance of cities, an issue that appear particularly significant within a knowledge base context such as that of SNB. Lastly, as it suggests some possible implications that a study on households sorting behaviours has for economic and spatial policies.

Consider e.g. highly-skilled educated households. If one can recover consistent information on the preference profile of this segment of the population, then he will be able to measure the marginal welfare effect that changes in the level of some amenities have on a group of households that is apparently very important for the SNB economy. Furthermore, if one is concerned about drawing skilled labour force into the region to boost the local economy, it is definitely sound to investigate what locational characteristics are considered more “attractive” by these workers. Last but not least, assuming that the local knowledge-based sector continues to expand in the future, it would be interesting to use our sorting model to see how an increase in the concentration of the skilled would affect the housing market and the locational equilibrium. As it will be shown in the rest of the paper, many of these issues can be effectively addressed through the analysis of household locational choice.

4. The locational choice of households

When people and firms can freely move across cities the spatial equilibrium condition ensures that in principle, for utility to be equal across space, higher nominal wages available in a location would be offset by a lower level of amenities or by higher housing cost. Within a city, however, wages are assumed to be fixed, as urban areas are meant to represent localized labour markets. Within cities, however, people can easily commute and work outside their neighbourhood, so individuals' earnings do not completely depend on their place of residence. All sorts of amenities become relatively accessible to local dwellers at the city level. Nevertheless, the place of residence still matters a lot when it comes to individuals' welfare. The areas of a city can still differ substantially under many aspects, so that people look very carefully at the amenities available at the very local level when they choose their dwelling. That said, a question that may arise when analysing the distribution of households across the urban space would be: what are people attracted to when they decide where to live? This is also what the present work is mainly concerned about. More precisely, we want to see how - and to what extent - certain amenities influence individuals' locational choice by investigating the process of households' sorting within a specific urban area.

Looking at households' sorting behaviour is of particular interest if one wants to study households'

preference for locational characteristics. Indeed, each location can be depicted as a composite commodity that allows consumers enjoy a specific bundle of amenities. As a rational household is expected to pick his preferred location out of the alternatives he has available, his choice will immediately reveal the researcher which location returns to that specific consumer the highest level of utility. Among other things, the Tiebout (1956) seminal model of sorting could be interpreted in these terms, i.e. as an attempt to demonstrate that, with perfect mobility and spatial variation in amenities, by “voting with their feet” people reveal their preferences for local amenities within a market of local public goods. In light of that, it is straightforward to see that the equilibrium outcome resulting from consumers' sorting will eventually reflect the distribution of households' preferences for locational characteristics.

Since households pay for enjoying location-specific amenities primarily through the price of housing, a viable way of studying households' preferences within a sorting framework is by looking at the housing market. Let households demand housing in a specific locations based on their preferences for local amenities. With fixed house supply, one would have that the demand for housing underlying households' sorting will naturally influence the house prices, so that these will adjust until the market clears. Therefore, the equilibrium will be characterized by a situation where, given the level of local amenities and house prices, everybody has chosen its utility-maximizing location and nobody wants to move. Again, as it is common in spatial economics, the variation in the price of housing will ensure that the relative attractiveness of each location will be capitalized by the corresponding property values, so that eventually utility is equalized across space.⁴

Hedonic price models have long represented a powerful tool for the analysis of heterogeneous goods such as housing. In this type of models, the price of a good is expressed explicitly as a function of its attributes⁵. This approach turns out to be extremely useful if applied on the housing market, especially if one considers that the value of local amenities such as public services, pollution, parks, etc. cannot be derived directly from market behaviours as there is no market for these “goods”. As demonstrated by Rosen (1974), indeed, when the price of a good is expressed as function of its quality, the partial derivative of this function with respect to its characteristics reveals the consumers' marginal willingness to pay for that characteristic. As the market clears, one can then make a regression analysis on the equilibrium property value so as to obtain an average welfare measure for marginal changes in the level of local amenities.

However, some authors have speculated that estimates based on hedonic price methods return only a rough measure of consumers' mean preferences for housing attributes, see e.g. Epple (1987) and

⁴ The monocentric city model is a typical example of how such equilibrium condition can be used to model locational choice behaviours and housing market dynamics at the city level. In this model, among the most important in urban economics, it is shown that, with homogeneous households, when locations differ among each other along a single dimension - i.e. distance from the (unique) employment centre - at every point in space this one-dimensional opportunity-cost will automatically be reflected on the corresponding price of housing. Thereby each location would result optimal, as the advantage of shorter commutes given by more “central” locations will systematically be offset by higher property values.

⁵ For a review of hedonic methods see e.g. Sheppard (1999).

Bayer et al. (2007). In a hedonic price setting, indeed, the equilibrium price is taken as given, and one has no information on the market conditions that concurred to determine such outcome. In case consumers have heterogeneous tastes for local amenities, for example, it would be hard to identify consistent structural parameters only on the base of the prevailing market prices, as the latter are shaped by the distribution of preferences (and all other market conditions) that exist at the time when prices are observed.⁶ Various answers have been proposed by scholars with respect to this type of identification problem (Kuminoff et al., 2010). Among others, the so called “equilibrium sorting models” have resulted particularly advantageous for the study of consumers' preferences in an equilibrium sorting framework. The following sections will introduce the main theoretical and econometric issues regarding this type of models.

4.1. The equilibrium sorting model

In equilibrium sorting models (ESM) the process of market sorting that characterizes the equilibrium conditions is thoroughly considered for the analysis of consumers' preferences. Differently than in more conventional methods such as the hedonic pricing, the market equilibrium is indeed described as the outcome of a process of consumers' sorting, and is not taken as given. This has important implications if considered e.g. within the context of an urban system. As it will be shown later, a structural model of locational choice can be embedded in the analysis, so that housing prices can be characterized as resulting from a general equilibrium model of locational choice. Because of that, ESM allow one to study more in detail how agents' sorting affects the market outcome. Moreover, one can use information on households' characteristics to describe the locational choice made by agents with heterogeneous taste for local amenities, and identify the distribution of preferences that underlies the prevailing sorting behaviours.

In ESM, the hedonic approach is still used for modelling differentiated goods, as rents are assumed to capitalize the values of the local amenities available to dwellers. But house prices are not expressed simply as a function of their characteristics. Rather, they represent the equilibrium outcome of a structural model of locational choice, so as to reflect the decision that households make when choosing in which area of the city to live. By doing so, the model identifies the properties of any market equilibrium, in that it accounts for the possibility that the collective behaviour influence the market outcome. This is likely to affect in the first place the price of housing, as explained before. But potentially also other amenities that are endogenously determined through the sorting process. It is well known, for instance, that phenomena of social interactions such as feedback effects, peer effects, voting, etc., can have a sizable impact on the socio-economic

⁶ Bayer et al. (2007) provide an intuitive example that helps understand this point. They imagine a situation in which households value a single, discrete characteristic that is having a view on a beautiful landscape (e.g. the Golden Gate Bridge). As only few houses in the market allow consumers enjoy this view, the hedonic price of the view would generally reflect the marginal willingness to pay of households who attach a relatively high value to the view (rather than of households with an average taste). If, on the other hand, the view was widely available in the market, the equilibrium price would arguably be lower, so that the hedonic pricing would return the marginal willingness to pay of someone much lower in the taste distribution.

development of local communities.⁷ After introducing some of the features that distinguish the ESM from other techniques of demand estimation, it is important to see how, within the general sorting framework just described, the individual process of locational choice is specified.

For what concerns the choice alternative, the urban landscape is normally characterized as a finite system of discrete locations (e.g. neighbourhoods, municipalities, etc.), where each of them provides a certain level of amenities. To capture the relationship between amenities and property value, housing is conventionally treated as a homogeneous good or numeraire – e.g. unit of housing services – that can be consumed at a constant location-specific price⁸.

Given this finite discrete choice set, rational consumers are assumed to choose, among all existing alternatives, the location that returns them the highest level of utility. Following this, ESM generally refer to the theoretical framework of discrete choice models. Bayer et al. (2004), in particular, have developed a model of equilibrium sorting - the so called horizontal sorting model - that exploit the discrete choice methods with random utility that were introduced by McFadden (1973, 1978). Individual choice and market demand are thus derived through a probabilistic approach, namely by means of a multinomial logit model with additive random utility function. As recognized by many, employing such probabilistic framework can be advantageous for various reasons.

First, with this methodology each alternative is assumed to be a substitute for all other alternative locations, so that the attractiveness of each location is weighed relatively to the attractiveness of all other locations. This is important when studying individual locational choice, as it well reflects the various trade-off that households usually encounter when deciding where to live. Indeed, some locations may result more attractive because of e.g. their proximity to the employment centres, but less attractive if this is associated to a higher cost of living (as postulated in the monocentric city model) or negative externalities that may come from near industrial plants (noise, pollution, etc). Similarly, some areas (e.g. those closer to the inner city) provide good accessibility to consumption activities (shops, restaurants, theatres, etc.), but at the cost of higher rents and a poorer natural landscape. In principle, thus, a large number of locational characteristics can be taken into consideration in ESM. Even more interesting, households are also allowed to differ in the way they value these characteristics.

A further important advantage offered by the model with random utility is indeed related to the fact that, by implementing a technique similar to that used in mixed logit models, also the parameters in the utility function can be treated as random variables. Practically, this means that the coefficients of the location characteristics considered can be expressed explicitly as function of observable

⁷ See e.g. Bayer et al. (2011), Epple and Romano (2010) and Bayer et al (2007). Using ESM, Bayer et al. (2007) have shown e.g. that when the taste for local amenities varies with households' sociodemographic characteristics such as race or income, the tendency of dwellers to self-segregate into neighbourhoods increases dramatically.

⁸ Under certain assumptions on the consumers' utility function, it is indeed possible to separate the value that properties get from its structural characteristics (floor space, number of bathrooms, bedrooms, etc.) from the value they derive from local amenities, see e.g. Kuminoff et al. (2010).

households' characteristics such as education level, marital status, presence of children, etc. This enables the researcher to estimate how different groups of households value local amenities, and thereby measure the partial equilibrium welfare effects of a change in the level of certain amenities. Furthermore, once the structure of preferences is identified, it is also possible to make different types of counterfactual analysis on the base of the general equilibrium framework. For example, one could see how a variation in the level of local amenities would influence the housing market, or study how a change in the population's sociodemographic characteristics affects the equilibrium outcomes.

Thanks to their flexibility, sorting models with random utility have been used by researchers for the study of different types of issues pertaining to urban economics and households' locational choice. Examples are the survey on the demand for education in the US made by Bayer et al. (2007), the Klaiber and Phaneuf's (2010) analysis of consumers' preferences for open space amenities, and the research on the effects that cultural heritage has on households' sorting of van Duijn and Rowendal (2013). In this paper we also make use of a horizontal sorting model to estimate households' marginal willingness to pay for various urban amenities. Before looking more in details at the locational characteristics we are concerned about, it is necessary to first provide a clearer description of the model used in this thesis. The following paragraph is thus intended to provide a formal specification of the econometric model on which our estimations are based.

4.2. Econometric specification of the residential sorting model

The econometric framework used in this paper is based on the sorting model with random utility developed by Bayer et al. (2004), and follows closely the methodology of van Duijn and Rouwendal (2013). Therefore, we define a Multinomial Logit Model where each household chooses its utility-maximizing location out of all alternatives available in the choice set. Let $N = \{1 \dots n\}$ be a set of n different alternative locations. A population of households, indexed by $i = \{1 \dots I\}$, is assumed to maximize a linear indirect utility function of the type:

$$V_{i,n} = \sum_{k=1}^K \alpha_{i,k} X_{k,n} + \varepsilon_{i,n} \quad (1)$$

This equation identifies the utility returned to household i by the choice alternative n . Utility is thus expressed as function of a set $K = \{1 \dots k\}$ of k different characteristics, where $X_{k,n}$ is the value of the k -th characteristic of alternative n , $\alpha_{i,k}$ the coefficient corresponding to the k -th characteristic, and $\varepsilon_{i,n}$ is the stochastic part of the utility function. More precisely, the ε 's reflect the i -th household-specific preference with respect to alternative n . For the model to be identified, these random terms are assumed to be independent, identically distributed (IID) extreme value type I distributed. Such assumption, which is basic in multinomial logit models, conditions individual relative preferences, in that it “restricts the substitution between the choice alternatives at the level of the individual actor” (van Duijn and Rouwendal, 2013). However, as recognized by many, if the heterogeneity among agents is pronounced this property is not likely to hold at the population level.

If that is the case, the model estimation would yield biased structural parameters.

To overcome this problem, Bayer et al. (2004) have exploited a technique similar to that used in mixed logit model. In order to incorporate individuals' heterogeneity into the model, they specify the coefficient $\alpha_{i,k}$ as function of households' characteristics. The i 's subscript that compares in the coefficients is indeed representing the fact that these α 's are individual-specific parameters. Formally, these are expressed as:

$$\alpha_{i,k} = \beta_{0,k} + \sum_{l=1}^L \beta_{k,l} (Z_{i,l} - \bar{Z}_l) \quad (2)$$

where L is the number of households' characteristics, $Z_{i,l}$ is the value of the l -th characteristic of household i , \bar{Z}_l is population mean of characteristic l , and the β 's are coefficients. Based on this specification, preference heterogeneity is identified by interacting observed households' characteristics (education level, age, etc.) with the set L of location-specific attributes. On the base of equation (2), the utility function can be rewritten as:

$$V_{i,n} = \sum_{k=1}^K \beta_{0,k} X_{k,n} + \sum_{k=1}^K \left(\sum_{l=1}^L \beta_{k,l} (Z_{i,l} - \bar{Z}_l) \right) X_{k,n} + \varepsilon_{i,n} \quad (3)$$

In this equation, the first term can be interpreted as the utility attached to alternative n by the average household, while the second term is the deviation from the average utility of household i , see e.g. van Duijn and Rouwendal (2013).

However, it is often the case that one or more of the characteristics that influence the locations' attractiveness are not included in the utility function, e.g. because these characteristics are not measurable or because they are not considered by the researcher. In order to account for such unobserved heterogeneity across locations, we include an additional term in equation (3):

$$V_{i,n} = \sum_{k=1}^K \beta_{0,k} X_{k,n} + \xi_n + \sum_{k=1}^K \left(\sum_{l=1}^L \beta_{k,l} (Z_{i,l} - \bar{Z}_l) \right) X_{k,n} + \varepsilon_{i,n} \quad (4)$$

The term ξ_n , which intends to capture all the unobservable qualities associated to location n , is assumed to be valued in the same way by all types of households. In principle, it is important to take these unobserved characteristics into consideration, as it is well known in econometrics that if it happens that something in the error term is correlated with any of the independent variables, the coefficients estimated for those variables will be biased. Potentially, this is a relevant issue in the present context, as it is expectable that one of our control variables - housing price - is correlated with unobserved characteristics that influence the attractiveness of the locations in the choice set. As anticipated earlier, not addressing such endogeneity issue would lead to a spurious estimation of the price coefficient derived with the sorting equilibrium. Specifically, consider the probability that household i chooses location n :

$$P_{i,n} = \frac{\exp(V_{i,n})}{\sum_{n=1}^N \exp(V_{i,n})} \quad (5)$$

The maximization of the probability that the individual choice is correctly predicted ensues from of a log-likelihood function of the form:

$$ll = \sum_i \sum_n Y_{i,n} \ln(P_{i,n}) \quad (6)$$

where $Y_{i,n}$ equals 1 if household i chooses alternative n and equals 0 otherwise. Given eq. (6), the predicted market demand for alternative n can be derived by aggregating the individual probability over the population of households:

$$D_n = \sum_i P_{i,n} \quad (7)$$

Hence, the market clearing conditions are identified by imposing Demand for housing to equal Supply:

$$D_n = S_n, \forall n \Rightarrow \sum_i P_{i,n} = S_n \quad (8)$$

Now, following the equilibrium condition in equation 5, relatively high prices might result for some locations because of a positive ξ (or relatively low prices if ξ is negative). However, as the price levels are not explainable on the basis of the observed characteristics, the researcher would be tented to conclude that housing prices have apparently a relatively low impact on households' locational choice. The reality would be instead that households actually respond to the unobserved quality ξ . In order to deal with the endogeneity of housing price, we trace the methodology introduced in Bayer et al. (2004). Therefore, we estimate the model on the base the two-step procedure (Barry et al., 1994, Bayer et al., 2004). To see how this works, we first rewrite equation 2.4 as:

$$V_{i,n} = \delta_n + \sum_{k=1}^K \left(\sum_{l=1}^L \beta_{k,l} (Z_{i,l} - \bar{Z}_l) \right) X_{k,n} + \varepsilon_{i,n} \quad (9)$$

$$\delta_n = \sum_{k=1}^K \beta_{0,k} * X_{k,n} + \xi_n \quad (10)$$

Based on equation 9, we first use a multinomial logit model to estimate the vector of the mean indirect utilities δ 's (i.e. the portion of utility that refers to the average household), together with the coefficients of the interaction terms $\beta_{k,l}$. Afterwards, in the second step, we treat the δ 's as dependent variables and we make simple OLS regression analysis based on eq. (10). Before describing how our instrumental strategy works, it is important to notice that the we do include the unobserved characteristics ξ in the δ 's, see equation (10). This has two relevant implications: first,

that in the first step the δ 's, functioning as a sort of alternative specific constants, ensure that the estimated coefficients ($\beta_{k,l}$) are such that the market equilibrium of equation (8) will always be fulfilled; second, that the unobserved characteristic ξ have an impact only at the aggregate level, as it affects only the mean indirect utilities. This implies that, as at the individual level prices and locational characteristics (both observable and unobservable) are taken as given, the coefficient of the cross effects estimated in the first stage will not be spurious (see e.g. Bayer et al., 2004, van Duijn and Rouwendal, 2013).

For what concerns the endogeneity of the housing price, we now know that this would affect the OLS parameters derived in the second-stage, see equation 10. We therefore proceed with a 2SLS following the instrumental variable approach of Bayer et al. (2004). Specifically, we instrument the housing prices with the vector of prices that would clear the market if there were no unobserved heterogeneity. In other words, by setting ξ_n equal to zero, we construct an instrument that is necessarily uncorrelated with ξ_n , and arguably also correlated with the observed housing prices. However, as the $\beta_{0,k}$'s are initially not known by the researcher, this is practically executed by mean of an iterative procedure, see e.g. van Duijn and Rouwendal, 2013.

With respect to the framework used in Bayer et al. (2004), and van Duijn and Rouwendal (2013) we introduce an additional term in our model of locational choice that is a component representing all types of costs a household incurs if he decides to move to a new location. The novelty here is that moving costs arise whenever households choose to move to a dwelling (n_t) different than the one he occupied in the previous time period (n_{t-1}). In other words, we assume that in each time period households decide whether to move to a new house or stay in his original dwelling. If he moves he will incur moving costs, and he will choose the location that yields the highest level of utility. Formally:

$$V_{i,n}^{MC} = V_{i,n} - \alpha_i^{MC} MC * I_{i,n} \quad \forall I_{i,n}: \begin{cases} = 1 & \text{if } n_t \neq n_{t-1} \\ = 0 & \text{otherwise} \end{cases} \quad (10)$$

A rational household is expected to change residence if the gain in utility he obtains at time (t) outweigh the costs of moving. It must be noticed that we exclude forward looking expectations made by households, so what we propose is a static model of locational choice with imperfect mobility of households. Furthermore, we assume that mobility costs are identical across all locations different than the one chosen in the previous period, and that they are function of observable household characteristics. We therefore express α_i^{MC} as:

$$\alpha_i^{MC} = \gamma_0 + \sum_{l=1}^L \gamma_l (Z_{i,l} - \bar{Z}_l) \quad (11)$$

This allows us to rewrite $V_{i,n}^{MC}$ as:

$$V_{i,n}^{MC} = \delta_n + \sum_{k=1}^K \left(\sum_{l=1}^L \beta_{k,l} (Z_{i,l} - \bar{Z}_l) \right) X_{k,n} + \left(\gamma_0 + \sum_{l=1}^L \gamma_l (Z_{i,l} - \bar{Z}_l) \right) MC * I_{i,n} + \varepsilon_{i,n} \quad (12)$$

Absence of transaction costs is a standard assumption when investigating consumers' behaviours and market performances, and it is explicitly conjectured also in baseline models of urban economics and agents sorting (Tiebout, 1959, Muth, 1969, Bayer et al., 2004). However, recent studies have attempted to demonstrate that mobility costs may in fact be relevant for households (Bayer et al., 2009, Caetano, 2010, Bayer et al., 2011, Epple et al., 2012). Epple et al. (2012), for instance, argue that mobility costs limit the possibility for households to change their residence so as to meet the various needs arising at various points during the life. Bayer et al. (2011) claim that when moving is costly households can hardly re-optimize location in every period, so that they will have to speculate closely on future circumstances before choosing their residence. Bayer et al. (2009), moreover, postulate that the difference in the level of amenities is not fully capitalized in property values if there are frictions in the housing market. Thereby, they demonstrate that mobility costs significantly affect the estimation of the marginal willingness to pay for locational characteristics.

We suspect as well that costs of various natures could restrain agents from making their optimal choice, so we relax the assumption of perfect households' mobility. It is indeed reasonable to expect that there are substantial costs households incur when moving to a new house. Apart from the financial expenses that generally arise when dealing with housing (transaction costs, realtor fees, etc.), it is important to consider also the broad range of non-monetary, psychological costs that relocating normally entails for people. Things like moving away from a familiar environment, the weakening old social relationships, and the effort requested to settle in a new place, are arguably strongly considered by households, and may have a relevant impact on their attitude towards the change of residence. That said, it is important to look more closely at the possible implications of moving costs with respect to the estimation of consumers' preferences. For instance, imagine a situation where households are somewhat discouraged from moving to their preferred location. If the analyst assumes that agents are freely mobile across locations, it would be straightforward for him to conclude that each household has chosen his utility-maximizing location. In reality, it might be that some households decided to stay in their original location because the benefit they could obtain from relocating was not as large as to outweigh the costs of moving. Consider for example the problem in equations (10)-(12). In period t household i decides to move from his original location n_{t-1} to a new location h if the utility he receives from this choice exceeds the utility that he receives from choosing to stay in its original location net of moving costs. That is when:

$$V_{i,h} - \alpha_i^{MC} MC > V_{i,n_{t-1}} \quad \forall h \neq n_{t-1} \quad (13)$$

To see how this could bias the estimation of the marginal willingness to pay, suppose that in period t a location becomes relatively more attractive because of some improvements in its physical or socio-economic characteristics. With perfect mobility one would expect that the demand for that

location would increase to some extent, so that a higher price would result from the prevailing locational equilibrium. If relocating is costly, instead, the value of that improvement would be discounted by households based on their moving costs. Because of that, the impact that the change in the amenity's level has on property values would eventually be smaller, so that the households' marginal willingness to pay estimated for that amenity would be biased downward.

In sum, overlooking moving costs could potentially bias our estimation of the alternative specific constants as well as of the other coefficients in the model, so we embed mobility costs into our model as shown in equation (10). In principle, if the assumption of perfect mobility holds one would expect that the results based on equation (1)-(4) are the same as those based on equation (10)-(12). Eventually, by comparing the two sets of coefficients, we could empirically test whether or not the assumption of households' perfect mobility holds. Moreover, by mean of the interaction terms in equation (12), we could also see whether the impact of moving costs vary among different types of households.

5.Data and Descriptives

Our analysis of households' location choice focuses on a specific Dutch region, that is the South-east North Brabant. Following the EUROSTAT territorial nomenclature, the SNB's area corresponds to one of the 40 NUTS-3 (COROP) regions in which the Netherlands is subdivided. Such spatial scale is quite suitable for our study, as it reasonably reflects a localized labour/housing market⁹. For what concerns the alternatives in the choice set, we consider as spatial units the 21 municipalities composing SNB, see Figure 2. Two cities are at the top of the regional hierarchy, i.e. Eindhoven, which with over 21,000 inhabitants is by far the most populous city, and Helmond, which has a population of approximately 90,000 people. For the rest, the regional spatial structure is similar to that of most Dutch regions, with a number of satellites cities surrounding the two main centres, and the least populated municipalities being those in the more peripheral areas of the region.

In order to estimate our residential sorting model we make use of data on both locational and households characteristics. For the latter, we refer to a survey – the Netherlands Housing Research 2012 (WoON) – carried out by the Dutch Interior Ministry and the Central Bureau of Statistics (CBS) every 4 years. WoON 2012 provides information on 69,339 Dutch households (1979 for the Eindhoven region), namely on household characteristics as well as on the conditions and location of their corresponding housing. To ensure representativeness of the sample for the national and local population, also household-specific weighting factors are included in the survey. These weights are used throughout all descriptive and econometric analyses made in our study.

⁹ For detailed information on the EUROSTAT “Nomenclature of Territorial Units for Statistics” (NUTS), see http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction. According to Brainport Eindhoven (2012), 83 percent of the local population commute within the regional boundaries, which is in line with the high rates of cross-regional commuting characterizing the Netherlands. This value suggests that the region can hardly be thought as a perfectly closed, circumscribed market. This implies that the assumption of having localized housing market should be considered more ideal than factual.

As shown in the previous section, we express the model's parameters as function of household socio-demographic characteristics so as to see how different types of households differ in their taste for local amenities. Given the knowledge intensive profile of the SNB economy, a great deal of attention is directed in this study to the preferences that more educated households have for a number of locational attributes. Beyond the level of education, there are other aspects that are important to take into account for explaining households' location choice.

Figure 2. 21 municipalities in South-East North Brabant (2012)



As shown in the previous section, we express the model's parameters as function of household sociodemographic characteristics so as to see how different types of households differ in their taste for local amenities. Given the knowledge intensive profile of the local economy, a great deal of attention is directed in this study to the preferences that the more skilled households (i.e. the college educated) have for a number locational attributes. Beyond the level of education, there are other aspects that are important to take into account for explaining households' location choice.

For example, it might be important to control for marital status, as singles and couples could have different constraints (budget availability, co-location problems, etc.), and different priorities with respect to municipality characteristics (housing, social atmosphere, etc.), see e.g. Costa and Kahn (2000). In order to capture the life-cycle dimension of households' preference, we consider also the age of the head of the households. This can be relevant with respect to preferences for social atmosphere, as well as to sensitivity to commuting and moving costs, especially for households that are close to retiring from their work. The presence of children can as well influence the residential decision of households. The need for space increases with the size of the household, so larger families could have a higher propensity to live in areas where rents are relatively cheaper. Moreover, households with children might also want to have good accessibility to services such as

kindergarten, schools, recreational areas, etc., that are typically concentrated in large cities. As level of educations, marital status, age and number of children are all provided in the WoON survey, we are able to control for these characteristics.

For estimating our residential sorting model, we use categorical variables to control for educational level, marital status and presence of children, while age of the head of the family is treated as a continuous variable. More precisely, we differentiate highly educated households – i.e. those with at least a college or university degree – from those with a lower level of education. With respect to marital status, we simply distinguish between singles and couples. The former are considered highly educated if both parents hold a university degree (power couples). Relatively to family size, we use a dummy that takes the value one if in the household there is at least a child under the age of 18. Table 1 reports descriptive statistics for each of these four households' characteristics. We can see for example that power couples together with power single add up to some 19 percent of the sample population. Thus, it results that about a fifth of the SNB population owns at least a college degree. We can expect that, if disaggregated for individuals, such share would get closer to the level found in other statistics pertaining to the spatial concentration of the higher educated (see e.g. Figure 1).

Table 1. Descriptive statistics household characteristics

Variable	Mean	Std. Dev.	Min	Max
Highly educated single	.0872	.282	0	1
Highly educated couple	.1055	.307	0	1
Couple	.589	.492	0	1
Household with children (-18)	.341	.474	0	1
Age head of the household	51.983	17.541	18	95

Source: WoON 2012. Number of observations is 1732.

As the analysis of moving costs is a central issue in this article, it is useful to investigate also the propensity to change dwelling observable in our sample of households. The WoOn 2012 survey provides information on the year when respondents moved to their 2012 house, and also on the municipality where respondents were living before moving to their 2012 residence. These pieces of information are crucial for our analysis of moving costs. On the one hand, having information on the settling year allows us to identify those households that changed location more recently. Within our static sorting model framework, these category of households – which we will call for simplicity “*movers*” – are those who were willing to incur in moving costs, i.e. those for whom the benefits from moving to their preferred location outweigh the costs of relocation (see equation 13). On the other hand, knowing the previous location of respondents enables us to recognize, within the choice set we consider for our estimations, which alternatives imply relocation for each specific household (see equation 10). With these two pieces of information available, we can easily customize the choice set to the different circumstances faced by the households in our sample. Eventually, if there is enough variation in the households' choice behaviour, it would be possible to analyse empirically how households' residential sorting is in fact influenced by mobility costs.

Table 2 reports descriptive statistics for movers, and for subcategories of movers. Movers are specified as households who move to their 2012 house later than the year 2007.¹⁰ We can see that about a quarter of the surveyed households changed residence in the 5 years before the survey was conducted. Therefore, although we refer to a relatively long time period, it seems that a large majority of households remained to their original location. This gives a first indication of the high propensity of people to stay in their original dwelling. Of course, we do not expect that all “stayers” would have actually preferred to move to a different house. What we suspect is instead that there is at least a part of them that did not move to the location that they preferred the most because of the various financial and psychological costs that this decision would imply.

Table 2. Descriptive statistics for movers and sub-categories of movers

Variable	Mean	Std. Dev.	Min	Max	No.
Move to different municipality	0.031	0.173	0	1	44
Move within original municipality	0.093	0.290	0	1	152
Move from outside SNB	0.028	0.164	0	1	48
Move but previous location unknown	0.114	0.318	0	1	186
Mover	0.265	0.442	0	1	430

Source: WoOn 2012. No. of observations 1732.

Furthermore, it is important to notice that some 40 percent of the movers did not specify in the survey the place where they were previously living. As this information is essential for the analysis of mobility costs, we cannot fully individuate the trade-off this group of households faced when choosing their residential location. Therefore, for the sake of our analysis, it becomes necessary to focus on the other types of movers, i.e. those for which such information is available. From the table it results e.g. that around 10 percent of the movers were previously living in a municipality outside the SNB region. We can expect that once these households decided to migrate to SNB, they were not influenced by mobility costs while deciding where to buy or rent a house. In other words, as migrants incur a certain amount of moving costs independently on which new location they choose, one can assume mobility costs do not influence their locational choice (at least within a static framework).

We can see also that only around 10 percent of the movers decided to relocate in a different municipality. It results indeed that a large share of movers – about 35 percent - did change dwelling, but stayed in their original municipality. It is reasonable to expect that such households opted for a different housing type, or for a neighbourhood that they considered relatively more advantageous. However, since municipalities are the spatial units we consider throughout our empirical analysis of

¹⁰ According to this arbitrary definition, we are indirectly assuming that our static model of households sorting applies to a 5 year period. This might be considered a relatively long period of time to be considered in a static setting. It is indeed reasonable to expect that households' location choice take place at shorter time intervals (e.g. every 2 years). Nevertheless, given the relatively limited sample size, we decide to refer to a longer time period so as to increase the number of observations corresponding to movers. However, it has to be notice that sensitivity analyses show that estimations do not change substantially with respect of the year to which we refer when constructing the dummy for movers.

households' location choice, it becomes quite cumbersome for us to embed this particular scenario in our residential sorting model. Following equation 10, indeed, we would assume that there are no moving costs when a household chooses a location (municipality) that is not different from the one they were residing before. At the same time, treating these movers as if they did not pay any type of moving costs sounds somewhat approximate, if not contradictory. Nonetheless, in order to follow our model in a consistent way, throughout our estimations we consider these households as a special case of movers. Namely, we treat households who moved within their original municipality's boundaries as if they chose not to change location, i.e. as non-movers. Said differently, we assume that households who remain in their original municipality do not incur moving costs. Although this may seem a stretching assumption, we believe that this is the most effective way to embed such special case of movers into our model of residential sorting.¹¹

For what concerns municipality characteristics, we consider:

- *Local Housing Market*, which we capture through a municipality-specific housing price index that corresponds to the price of one unite of a standard house (housing service). This index is computed by mean of standard hedonic price estimation with municipality fixed effect. On the base of this index, the most expensive areas (above €210,000) result those located South, South-East of Eindhoven city (Valkenswaard, Waalre, Bergijk) together with other locations surrounding Eindhoven (Oirshot, Veldhoven, Son&Bruegel). Compared to the other municipalities, housing result not particularly expensive in the two main centres. Eindhoven presents an average level of prices (€205,147), while Helmond stands as the second cheapest location (€173,075).
- *Natural Amenities*, which we identify through both natural and water coverage (Km²) in the municipal territory. These are generally more abundant in less build-up and less populated areas such as those at the edge of the region. The idea is that households might value accessibility to open space as well as living in an area with a beautiful natural landscape.
- *Urban Amenities*, identified by various services and facilities that households enjoy “consuming” during their leisure time. As substantial scale economies exist in the production of such “goods”, a large amount of shops, bars, cultural and sport facilities is typically found in more populous municipalities. Given the closed link between the level of urban amenities and the municipal population, there is a high correlation in the level of the various types of consumption available at the municipality level (see Table 3). Hence, in our estimation we consider as a locational characteristic only the local number of catering services (i.e. shops, bars, restaurant, etc.). This

¹¹ However, it has to be notice that the psychological costs of moving, i.e. the most substantial type of costs related to relocation (e.g. loss of social contacts, no familiarity with the local environment, etc.) might be trivial if one continues to live in the same “area”. In a way, the fact that households might be inclined to avoid the psychological costs arising with long-distance migration makes our assumption less unfounded. That said, treating them in the same way as we treat those who changed municipality might also be inappropriate, as the amount of moving costs they incur is arguably different between these two categories, given the low impact that the psychological costs have on those who remain in their original municipality. On the other hand, treating the special case of movers as indifferent with respect to moving costs would be unreasonable, as they in fact weighed the various alternatives differently.

variable is thus meant to function as a proxy for the overall local level of urban amenities. By doing so we avoid multicollinearity-type of problems that would otherwise arise if many of such facilities were considered simultaneously. A simple univariate regression between number of catering companies and municipal population (see Figure 3 in Appendix A) suggests that a strong relationship between amount of consumption amenities and city size exists also at the municipality-level.

- *Accessibility*, interpreted in various ways. First, we include as a locational characteristic the distance between the municipality (core) and the nearest 100.000 jobs, so as to have a rough measure for the relative accessibility to the regional employment concentrations. Moreover, we also take into account the accessibility to various modes of transport by including distance to the nearest intercity station and distance to the nearest motorway ramp. These values are computed considering the municipality geographical core. In principle, a good accessibility to transport infrastructures can reduce overall travel times, so that households can reach more easily not only their work place but also other amenities that are located in their municipalities as well as in other areas of the region. Moreover, it has to be notice that intercity stations exist only in Eindhoven and Helmond. Since the highest amount of jobs and shops are found in these two municipalities, one can expect that distance from the nearest intercity station represents also a good approximation of the municipalities' relative distance from the main regional concentrations employment and services.

Table 3. Correlation matrix number of catering companies and other leisure/consumption urban amenities

	Museums	Theatres	Cinemas	Catering Comp.	Sport Facilities
Museums	1				
Theatres	0.69	1			
Cinemas	0.75	0.71	1		
Catering Comp.	0.84	0.74	0.87	1	
Sport Facilities	0.82	0.73	0.77	0.94	1

Source: ABF(2007)

Table 4 reports descriptive statistics for the locational characteristics that we take into account throughout our econometric analyses. In order to get the picture of the spatial structure of South-east Brabant region, it is useful to briefly go through some municipal-level descriptive statistics. For example, it might be interesting to see if a distinctive locational pattern arises when considering the distribution of skilled households across the region. Figure 4 (Appendix A) reports the concentration of highly-educated households for each of the 21 municipality. From the figure it seems that highly-educated households are relatively more concentrated in the medium-sized municipalities surrounding Eindhoven. In particular, high concentrations are found in municipalities that are both close to the main centres, and have relatively low housing price¹², a sort of compromise between short commutes, which is important given their relatively higher value of time, and their preference for spacious dwellings.

¹² Among the top 7 municipalities for concentration of highly-educated households, only 2 are also among the 7 most expensive locations (price of a standard house).

Another important factor to keep in mind is that the quality of housing can differ substantially across locations, so that the spatial distribution of households could in part reflect the preference that households have for certain type of housing. Detached or semidetached houses, for instance, might represent a good option for affluent families, not only because they can afford them better than others, but also because they like having a spacious dwellings, garden, parking space, etc. We can expect that households' preferences develop during life-cycle. Figure 5 Appendix A describes the demographic composition of the population of the 21 municipalities. One can speculate e.g. that elderly households prefer to live in comfortable and quite dwellings, typically more abundant in the outskirts, while singles and younger people are usually not so reluctant to live in apartment buildings such as those largely available in cities like Eindhoven and Helmond. Unfortunately, we do not have data on the municipality housing stock, so we cannot deal with this issue in a proper way. We can expect, however, that detached and semidetached houses are relatively more abundant and affordable out of the main urban centres. This might be important especially when considering the impact of age and level of education on households' location choice.

Table 4. Descriptive statistics locational charac

Variable	Data source	Mean	Std. Dev.	Min	Max
Price of standard house (Euros)	NVM(2007)	200235.3	12891.48	171868.4	217833.1
Natural Coverage (Km2)	CBS(2007)	24.569	9.305	10.07	49.34
Water coverage (Km2)	CBS(2007)	1.312	.655	.37	2.658
Num. of catering companies	ABF(2007)	97.667	144.837	31	709
Distance to the nearest 100.000 jobs (Km)	ABF(2007)	13.515	4.994	4	23.383
Distance to the nearest intercity station (Km)	ABF(2005)	10.372	5.154	2.491	22.641
Distance to the nearest highway ramp (Km)	ABF(2000)	4.464	3.184	1.166	12.05

6. Estimation Results

This section reports the estimation results for our model of residential sorting. As described in section 4, we implement the two-stage procedure developed by Bayer et al. (2004). We first present the results obtained in the first-stage regression, showing the impact that the introduction of the moving cost component has on the estimates of the Multinomial Logit Model. Subsequently, we briefly go through the second step procedure, where the 2SLS regression is implemented. Finally, we present the structure of households' MWTP that we compute by using the parameters derived in the two aforementioned steps of estimation.

6.1. Results from the first-stage regression of the residential sorting model

In the first step of our estimation procedure we follow the methodology used in Bayer et al. (2004) and van Duijn and Rouwendal (2013) to derive the parameters that maximize the probability that households' location choices are correctly predicted (equation 5-6). Therefore, we run a

Multinomial Logit Model where we consider the 21 SNB's municipalities as spatial units. Specifically, we estimate a vector of alternative specific constants, together with a set of coefficients involving the interaction between household and locational characteristics, see equation 9. The vector of mean indirect utilities will tell us the utility attached to each municipality by the average household. The coefficients obtained for the cross effects will instead reflect the change in the indirect utility estimated for a marginal deviation of the household characteristics from the corresponding sample average.

For deriving the distribution of households' preferences we will refer to the owner-occupied sector only. We do include observations for the rental sector (about 40 percent of the sample), as they increase the model's overall representativeness of the population. But we try to detach the two sectors throughout the estimation procedure. The reason is that in the Netherlands a large part of the rental market is subject to a series of rigidities (rent control, waiting lists, etc.) which distort the market and arguably also the process of households' sorting. Therefore, similarly to van Duijn and Rouwendal (2013), we include the rental sector in each municipality as an alternative for the owner-occupied sector, and we estimate two separate sets of coefficients, i.e. one for the owner-occupied and one for the rental market. This implies having 42 choice alternatives in total, i.e. two for each municipality. Table 4¹³ reports the coefficients for the interaction terms between households and locational characteristics obtained in the first stage estimation. The sign of the coefficients estimated reveals the relative sensitiveness to the level of locational characteristics of the various categories of households. For example, if a positive coefficient results for the household type H with respect to the locational characteristic A, this would mean that a marginal positive change in the level of A has a larger impact on the utility of H than on utility of the average household in the sample. As for every other table of results provided in the paper, we report only the set of coefficients corresponding to homeowners.

The two columns in Table 5 refer to two different estimation scenarios. Respectively, column 1 reports the estimates obtained when moving costs are not included in the model (equation 9), while column 2 reports the estimates obtained by including moving costs (equation 12). It is important to notice that we address moving costs only in the first-step regression. Mobility costs are indeed individual-specific. As they have nothing to do with the average preference that we analyze in the second stage of our estimation procedure, they have to be considered only when dealing with the individual choice problem, i.e. when estimating the Multinomial Logit Model. Before commenting on the cross coefficients found for the locational characteristics, it is important to provide a thorough explanation of the effects produced on the estimates results by the introduction of moving costs.

¹³ It has to be noticed that for the estimations of the first-stage parameters, along with the 7 municipality characteristics mentioned in section 5, we also include a dummy variable for Eindhoven and Helmond. This can be useful especially for the analysis of preferences for consumption amenities. Given that most catering services are concentrated in these two municipalities, this sort of municipality fixed-effects are expected to reduce the collinearity that might exist between the amount of consumption amenities and other uncontrolled municipality characteristics that could be relevant particularly with respect to the two larger cities (e.g. crime, congestion, locational image and social atmosphere, etc.)

Table 5. First stage results for homeowners: estimated parameters for interaction terms and moving costs

Municipality characteristics	Household characteristics								
	Mean		Highly educated households		Age (+10 years)		Households with children (-18)		Couples
Standardized house price (€)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
No. of shops		5.674 (-0.258)***	8.630 (-0.578)***	0.122 (-0.00749)***	0.0476 (-0.018)***	4.717 (-0.229)***	5.570 (-0.551)***	1.124 (-0.259)***	-5.727 (-0.636)***
Nature (km ²)		-0.00221 (-0.0000683)***	-0.0641 (-0.00017)***	-0.000062 (-0.000002)***	-0.000245 (-0.00000537)***	-0.000951 (-0.000063)***	0.000956 (-0.000167)***	-0.000831 (-0.000069)***	-0.0051 (-0.000191)***
Water(km ²)		0.0015 (-0.000095)	-0.0445 (-0.00214)***	-0.000053 (-0.0000297)	-0.00096 (-0.0000631)***	-0.00136 (-0.000937)	-0.0058 (-0.00201)***	-0.0323 (-0.00098)***	-0.0129 (-0.00226)***
Distance to 100,000 jobs (km)		-0.232 (-0.0126)***	0.523 (-0.0284)***	-0.00625 (-0.00038)***	-0.0278 (-0.000946)***	-0.0105 (-0.0115)	0.302 (-0.0269)***	0.0104 (-0.0129)	-0.138 (-0.0326)***
Distance to intercity station (km)		0.0605 (-0.00615)***	0.0215 (-0.0134)	0.000646 (-0.000176)***	0.00187 (-0.000428)***	0.0375 (-0.00536)***	0.122 (-0.013)***	0.0501 (-0.00614)***	-0.31 (-0.0152)***
Distance to motorway ramp (km)		-0.142 (-0.00602)***	-0.0797 (-0.0119)***	-0.00164 (-0.000155)***	-0.00233 (-0.00037)***	-0.0497 (-0.00469)***	-0.0696 (-0.0109)***	-0.0277 (-0.00545)***	0.215 (-0.013)***
Moving costs homeowner		-0.123 (-0.00386)***	-0.0593 (-0.00759)***	-0.00216 (-0.000102)***	-0.00819 (-0.00025)***	0.0116 (-0.00307)***	0.0581 (-0.0072)***	-0.0254 (-0.00359)***	0.0626 (-0.00841)***
Moving costs renter		-8.355 (-0.0247)***	0.616 (-0.0402)***	-0.0921 (-0.0014)***	-1.543 (-0.0419)***				0.248 (-0.0417)***
		-7.125 (-0.0298)***							

Variables are normalized to have mean zero, so the coefficients represent the estimated deviation from the mean indirect utility. Column 1 reports estimates for the interaction parameters of residential sorting model without moving costs (equation 9). Column 2 reports estimates for the interaction parameters of residential sorting model with moving costs, together with the parameters corresponding to the moving costs component. Standard errors in parenthesis. Dummies for Eindhoven and Helmond are included in both estimations. 90, 95 and 99% significance level are indicated, respectively, as *, ** and ***. Movers are identified by households who moved to their 2012 location later than the year 2007. Movers who stayed in their previous municipality of residence are assumed not to incur in moving costs. Movers who migrate from outside the SNB region, or whose previous location is unknown, are considered indifferent with respect to moving costs.

It is a good idea to start by comparing the results derived under various estimation scenarios. That is why we provide also the estimates derived by including moving costs with a single coefficient, i.e. not interacted with household characteristics. These are reported in Table A1. We can see that the introduction of the moving cost component with a single parameter produce a sizable change in the various estimates. This already constitutes a tangible piece of evidence about the impact of moving costs on the model estimations. As expectable, a negative and significant parameter result for single moving cost parameter. This parameter can be interpreted as the mean impact on indirect utility of moving to a different location, in our case to a different municipality. We report the mean moving cost effect both for homeowners and for renters. We find that on average relocation has an impact on renters that is relatively lower than on homeowners. This is consistent with the fact that renting is mostly a temporary solution for people, and that it entails more open and flexible transactions. On the other hand, buying a new house requires longer and more expensive transactions with sellers, estate agents, etc., which can result demanding especially for certain categories of households. A further explanation could be that owner-occupied housing represents a durable good on which dwellers “invest” money and time, so that they are more reluctant to move away once they have spent a good part of their life there, see e.g. Rouwendal, (2009).

It must be noticed however that larger and wealthier families typically own their house, while renting is chosen mainly by single and young households, or individuals who lack financial resources and wait for better housing and economic conditions. As presence of children, age and marital status are correlated with both housing-tenure and mobility, it becomes particularly interesting to add the interaction parameters to the single (constant) coefficient. The cross effects of moving costs can be found in column 2 of Table 5. The difference between the sets of estimates in column 1 and column 2 reveals once again that moving costs influence the estimation results. Moreover, it is important to notice that the impact of moving costs vary significantly across households. It seems indeed that factors such as age and presence of children make households relatively less inclined to move from their original residence. The low propensity of elderly households to leave their house is well-recognized in the literature and strongly related to the fact that psychological and settling-costs increases with age see e.g. Rouwendal and Thomese, 2010. It is reasonable to expect that moving becomes more prohibitive also when there are under-18 children living the household, since it could imply school disruption, loose of friends and other negative circumstances for the children. As mobility is particularly higher among single households, the positive estimate found for couples is hard to explain. Nevertheless, it could be that couples can better realize the change since they move together, while singles perceive more the costs of moving because they rely more heavily on the local social network. Finally, it seems that moving has a relatively lower impact on the higher educated. This might be due to the fact that wealthier households can afford more easily the financial costs of moving. Given their higher budget, they might also have access to a larger set of alternative, so that the possibility of improvement is somewhat higher for them.

To summarize, the comparison among the three sets of parameters suggests that the impact of moving costs is not homogeneous across households, and that assuming perfect mobility may lead to a spurious estimation of the interaction parameters. Apart from that, overlooking mobility costs is likely to affect also the estimation of the vector of mean indirect utilities. If that was the case, it would not be possible to recover consistent estimates of the average preferences of households, so that the resulting structure of MWTP would be biased and not reliable.

Table A.2 reports the mean indirect utilities relative to the owner-occupied market. These represent a measure of the average attractiveness of the SNB municipalities relatively to the alternative that we choose as reference, i.e. Helmond. At this point of the analysis, looking at the statistical relationship among the sets of parameters could effectively reveal the relative impact that the introduction of moving costs on the estimates. Reasonably, indeed, in case the impact of moving costs is trivial, one would find that there is a high correlation across the various sets of parameters. Therefore, we compute correlation matrices considering three different specifications of our residential sorting model. These are reported in Table A.3.a.-A.3.c. From the table it seems that the introduction of moving costs influences sharply our results. The correlation between the whole set of parameters respectively of the model without moving costs, and of the two models with moving costs, is about zero. Looking at the correlation between the vectors of mean indirect utility, the value is higher (0,77). As expectable, the variation is minimal when the model's specification with a fixed moving cost coefficient is confronted with the model with the interaction term (0,98). This is not surprising, as the introduction of the cross effects for moving costs aims mainly at revealing how the moving cost effect vary across households' types. Overall, however, the results evidence that embedding mobility costs in the analysis has important consequences on the model. We can then conclude that overlooking imperfect mobility of households could seriously undermine the reliability of our estimations of the MWTP. Our preferred estimates will therefore be those in column 2.

The parameters of the cross effects provide a first indication of how heterogeneous households differ in their taste for locational characteristics. For example, we can see that highly educated households, couples and households with children are relatively less sensitive to housing cost than the average household in the sample, while the opposite holds for elderly households. This is somehow consistent with the differences in income level that might exist among the various categories of households.¹⁴ Also the taste for accessibility to employment may be indirectly influenced by relative household income levels. Indeed, albeit wealthier households value strongly short commutes, they might result relatively less sensitive to employment accessibility because they can rely more on faster (and more expensive) transport modes (car), see Glaeser and Kahn, 2004. However, the higher value of time that normally characterizes higher educated (wealthier) households seems to play an important role when it comes to accessibility to transport infrastructures (railway and highway network).

The taste for natural amenities results quite heterogeneous among categories. Interestingly, there is some evidence that being a couple and the presence of children have a positive impact on the

preferences respectively for natural and water coverage. Finally, preferences for consumption amenities appear relatively stronger for younger and single households and also for those with children. Given the children-specific type of needs, the former might value having a large variety of public and recreational services (schools, sport facilities, etc.) close by. Within a relatively small region, however, living in the main urban centers may not be a primary condition for enjoying consumption amenities. Indeed, it is relatively easy for households' to reach consumption facilities available in nearby municipalities. Hence, as it will be explained better at the end of the section, spatial dependence between municipalities becomes particularly relevant for the analysis and the interpretation of taste for consumption amenities resulting for certain type of households.

6.2. Results from the second-stage regression of the residential sorting model

Following the first-stage estimates, we consider the vector of mean indirect utilities as dependent variables and we regress them on locational characteristics, see equation 10. The related parameters will reflect the impact that locational characteristics have on the indirect utility of the average household. Unfortunately, estimates based on the alternative specific constants obtained in the first-stage regression described above resulted not very useful for the analysis of the mean preferences. As expectable, indeed, the number of observations (21 municipalities) is too small to carry out consistent econometric analyses. Particularly problematic is that, despite we are able to derive a strong instrument for housing price, the 2SLS procedure does not converge even after many iterations. As the vector of average parameters is crucial for the computation of households' willingness to pay, this poses a serious threat to our empirical analysis. Given that, we decided to proceed by applying our model of residential sorting to a larger sample. Specifically, we estimate a similar residential sorting model but instead of focusing on the Eindhoven area we considered a much wider number of Dutch municipalities (399).¹⁴ By increasing the number of observations we could indeed rely on a larger variation in the alternative specific constant as well as in municipality characteristics, so that we could effectively perform the second-stage estimations.

As described in section 4, the 2SLS estimation technique we follow is the one introduced in Bayer et al. (2004). Price instruments are thus constructed by deriving the vector of prices that would clear the market in the absence of unobserved heterogeneity. Apart from housing prices, endogeneity is expected to be an issue also with respect to the local level of consumption amenities. The problem could be indeed that unobserved characteristics that influence the attractiveness of municipalities can also have an impact on the amount of shops, catering services, etc. located within the municipal boundaries. As we suspect that the number of shops is correlated with unobserved municipality

¹⁴ It has to be noticed that the residential sorting model estimated with data on 399 municipalities consider the same household and municipality characteristics we accounted for when analyzing the SNB region. When dealing with the larger sample, however, neither we control for moving costs nor we include any municipality fixed-effect. Hence, the fact that we combine sets of coefficients based on different estimation procedure for deriving household MWTP might undermine the validity of our welfare measures. If e.g. it is the case that moving costs have an impact on the estimation of the alternative specific constants, such approach risks being inconsistent. However, we presume that moving costs do not have a strong impact on the mean indirect utilities computed for the 399 municipalities. We thus believe that the marginal willingness to pay obtained with our strategy is sufficiently reliable.

characteristics, we also instrument our proxy for local consumption amenities in the 2SLS regression. The instrument we use in this case is an historical instrument, i.e. the 1830 municipal population level. The idea is that, given the straightforward relationship between local market size and amount of shops, and the persistency that is normally found in the population level, 1830 population is supposed to be a relevant instrument for local amount of shops. At the same time, one can expect that such instrument is also exogenous, since it is not likely that factors that influence locational attractiveness in the present time are linked to the factors that determined population level in 1830.

Table A.4. reports the results derived in the second-stage procedure. These parameters reveal the average impact that various locational characteristics have on municipality attractiveness. Column 1 refers to the simple OLS regression, while the other columns refer to the estimates obtained through the 2SLS procedure, that is: the parameters estimated when only housing prices are instrumented (column 2), and the estimates when both housing prices and shops are instrumented (column 3). By comparing the OLS with the 2SLS sets of parameters, we can see that the IV technique leads to a large variation in the estimates. The strongest change is related to the price instrument, as evidenced by the difference in the estimates between column 1 and 2. Indeed, the coefficient of housing price gets to a much larger negative value. This is in line with other researches dealing with residential sorting models (see e.g. van Duijn and Rouwendal, 2013).

Also the coefficients of the other municipality characteristics vary significantly when housing price is instrumented. Noticeably, the parameter related to number of shops becomes smaller and insignificant. Moreover, we can see that the overall picture changes very slightly when also shops are instrumented. This happens albeit various tests based on the 2SLS regression demonstrate that both instruments we use are strong and exogenous.¹⁵ Therefore, based on these results it seems that, even after controlling for the endogeneity of number of shops, there is a weak and inconsistent causal relationship between consumption amenities and average indirect utility. As it is generally argued that consumption amenities are increasingly important for households' residential sorting, this finding has important implications for our study. As anticipated earlier, however, considering the level of consumption services available in nearby municipalities could be helpful to understand the role that accessibility to consumption amenities has on households' location choice.

6.3. Marginal Willingness to Pay

Once derived the second-stage results, we can combine these estimates with the interaction coefficients obtained in the first stage so as to identify the structure of households marginal willingness to pay (MWTP) for locational attributes. Considering the four household characteristics

¹⁵ Regardless the results provided by the identification tests, some may question the exogeneity of the historical instrument by considering the impact that current population level could have on municipality attractiveness. Indeed, it might be that households care about the local population size when deciding their residential location. If that was the case, current population would represent an unobserved municipality characteristic. Given the correlation between 1830 and 2012 population level, the validity of our historical instrument would be jeopardized .

we accounted for to specify the coefficients of the indirect utility function (equation 2-3), and taking as reference the average price of a standard house in the SNB region, it is relatively easy to compute the MWTP that various groups of households have for municipality characteristics, see Appendix E. Table 6 reports the figure of the MWTP of the average household as well as of the other types of household. These are computed on the base of our preferred sets of parameters, that is the crosseffects derived when controlling for moving costs and the average effects on households' utility derived through the 2SLS procedure. We report the mean MWTP value in column 1, and the deviation from the mean in columns 2 through 5. Therefore, we provide a vector of mean willingness to pay (in terms of housing price in the municipality) for a marginal variation in the municipality characteristics, and the deviation from that mean corresponding to the different groups of households.

For what concerns natural amenities, we can see that the average MWTP for an additional square kilometer of natural coverage in the municipality of residence is €837, and €371 for an additional square kilometer of water coverage. Households with children seem to have a relatively strong preference for such amenities, as their MWTP for nature is about 25% higher than the mean, and more than three times higher for water coverage. It is hard instead to draw conclusions on the taste that the other categories of household have for natural amenities, as their MWTP for nature and water diverge.

Table 6. Marginal willingness to pay for municipality characteristics and for avoid moving

	(1) Mean	(2) Highly educated households	(3) Age (+ 10 years)	(4) Households with children (-18)	(5) Couples
Number of shops (+1 unit)	13	-118.9	-33.5	15.5	-26.1
Nature (+km ²)	837.2	-171.0	-105.5	205.3	-175.7
Water (+km ²)	371.4	991.9	743.9	1163.8	610.9
Distance to 100,000 jobs (-km)	3894.8	2978.1	-131.5	-162.2	946.9
Distance to intercity station (-km)	334.8 (ns)	1916.1	333.6	925.8	-1069.5
Distance to motorway ramp (-km)	880.8	1978.9	1165.5	-394.9	-418.5
Avoid moving to different municipality	112181.1	85844.6	54742.4	16471.6	-16478.7

Values are expressed in Euros. (ns) indicates that the value is not significant at the 10% level. Column 1 reports the mean MWTP of a marginal positive (+) or negative (-) change in the municipality characteristic, as well as of staying in the original municipality of residence. Columns 2 through 5 report the deviation from the mean MWTP for that type of household. Values and significance level of the mean MWTP and the corresponding deviations refer to the estimates reported respectively in column 3 table A.4. and column 2 of table 5.

Accessibility to employment and to transport infrastructures are calculated as the distance from the related municipality geographical cores, so it is puzzling to identify a precise monetary value that such characteristics have for the various groups of households. However, we can see that higher educated households present the highest MWTP for all three accessibility measures we consider. As

anticipated earlier, this finding is reasonably related to the relationship existing respectively between level of education, income, and value of time. We can see also that elderly households, i.e. those that are close to retire from work, are relatively less concerned about commuting distances. Co-location problem, on the other hand, is arguably the factor that can better explain the high MWTP found for couples (see e.g. Costa and Kahn, 2000). The latter, at the same time, value proximity to intercity station and highway ramp respectively 3 times and half time less than the average household. A reasonable explanation could be that singles and couples have different attitude towards leisure and social life, so that marital status produces a strong deviation from the mean preference for transport mobility and distance to the intercity stations (i.e. to the larger centers). Finally, we can see that households are WTP €130 more for a standard house if the number of catering services in the municipalities increases by 10 units. Households with children have a MWTP for number of shops that is about twice as much as the average, while the other three categories of households have all a negative and significant MWTP for living in municipalities with many consumption amenities.

In sum, the results confirm that there is a high degree of heterogeneity in the way households value locational characteristics. We find e.g. that highly educated households like having good access to jobs and transport facilities, but they value quite negatively places where consumption-type of amenities are largely available. Younger and single households, together with households with children, are those that seem to enjoy more from living in large municipalities. We expect singles to search for large marriage markets and social context. For households with children it might be that they have to cope with individual needs of different nature, so that they find more convenient to reside in locations where a larger variety of facilities and employment possibilities are available. On the other hand, elderly households “dislike” municipalities where consumption services are located. They instead prefer to live away from the employment concentrations, and especially close to the highway infrastructure. It results also that couples have very strong taste for places providing good accessibility, and that households prefer to “grow up” their children where both open space and consumption amenities are abundant.

It is important to emphasize also the fact that the estimated deviations of the mean MWTP hereby provided are based on observations referring to the SNB region only. That said, it is straightforward to expect that with a small spatial scale such as that of SNB inter-municipal spatial dependence is likely to be a relevant issue. This also on the base of the evidences provided by other studies that demonstrated how the amenities available in nearby municipalities influence households' locational choice quite consistently (van Duijn and Rouwendal, 2013). Controlling for the level of amenities available in neighboring municipalities would definitely help understand the role played by consumption as well as by the other amenities. For instance, one can presume that wealthier, highly educated households prefer to live in the satellites cities surrounding Eindhoven and Helmond because of the congestion and other urban disamenities that usually characterize these places, or because their preferred housing types are relatively more available outside large cities. At the same time, given that the distance between satellites municipalities and Eindhoven is quite short, they could easily reach the larger cities for accommodating their consumptive, recreational, and social

needs. Based on this, we can expect that the high MWTP we found for accessibility to transport facilities reflect the demand that certain groups of households have for accessibility to amenities available in other municipalities.

Similarly to how we did for locational characteristics, we can use the parameters estimated for the housing price and for moving costs to derive a monetary measure of households' willingness to relocate. We find that on average, in terms of housing price, a household is willing to pay about 110000 Euros to avoid moving to a different municipality. This amount appears extremely high, especially if we consider that the average price of a standard house in SNB is about 200000 Euros. Our finding suggests indeed that a owner-occupying household would bribe for housing an amount of money that is around 60 percent above its actual market value if that allows him to stay in his original location. Also relatively to the other MWTP this value seem high: the welfare effect of relocation is comparable indeed to a 130 km² increase in local natural coverage, and to the opening of 8600 new shops.

Nevertheless, it is important to point out that the way we model moving costs could lead to overestimating the moving cost effect. Indeed, we intend moving as a relocation to a (random) different municipality, and we assume that moving within the boundaries of the original municipality is equivalent to not moving. This implies that almost 10 percent of the households in the sample were considered as stayers while in fact they did result willing to change dwelling. Given that, we can expect that the adversity towards relocation is somewhat exaggerate when intramunicipality movers are treated as stayers. Said differently, the observed adversity towards relocation is likely to increase if one consider only "long-distance" movements, and exclude intramunicipality type of relocations. Furthermore, Bayer et al. (2010) highlight also that with forward-looking behaviors households are likely to discount also the stream of utilities corresponding to future time periods. This could also produce an upward bias in our MWTP estimate, as it could be that the psychological costs of moving are realized by households over multiple periods of time.

We can see also the MWTP to avoid moving increases with education level, age and presence of children. These types of households are found to be WTP respectively 75, 50 and 15 more than the average households to remain in their original location. This seems to confirm our idea that wealthier households are more reluctant to leave their house, as they tend to invest more in their dwelling. Finally, the fact that presence of children and age has a positive impact on the MWTP to stay in the original house confirms standard economic conjectures; see e.g. Rouwendal and Thomese, 2010, and Epple et al. Sieg, 2012.

7. Conclusions

In this thesis we conduct an analysis on residential sorting behaviors in the Eindhoven area to see how moving costs and various locational amenities weigh on the locational choice of households. Considering the 21 SNB's municipalities as spatial units, we perform an empirical analysis by

employing the horizontal sorting model developed by Bayer et al. (2004). We first use a discrete choice model to estimate the mean relative attractiveness of each alternative and a set of interaction parameters between household and municipality characteristics. Next, we make a 2SLS regression analysis on the vector of mean indirect utilities derived in the first-step to estimate the mean impact that municipality characteristics have on agents' utility. These two sets of estimates are eventually combined for estimating the MWTP that the different groups of household have for municipality characteristics.

Most importantly, we extend the standard residential sorting model by relaxing the assumption of perfect mobility of agents. A mobility cost component is thus added to the indirect utility function, so as to reflect the monetary and psychological costs households incur when they decide to relocate to a different dwelling. We find that moving has a strongly negative effect on households' utility. The study evidences indeed that households are considerably restrained from choosing a different location. Moreover, the first step estimates show that the introduction of mobility costs produces a substantial variation in all other parameters of the model. Therefore, we provide consistent evidences on the fact that overlooking imperfect mobility of households lead to biased estimates of the model's parameters.

Following the large coefficients estimated in the first step regression, the welfare effect that moving entails is substantial. We find that households would be willing to pay around 60% of the price of a standard house in order to avoid moving to a different municipality. It results also that the MWTP to stay in the original location varies considerably among the various types of households we consider. Highly educated households, for instance, are WTP 75% more than the average household for avoiding relocation. It seems that wealthier households are particularly reluctant to leave their house, perhaps because they generally invest more on their dwelling. This might be the factor that makes relocation so costly for the higher educated. Presence of children and age has also a strong impact on the level of the MWTP. Since the psychological costs of moving are expected to increase with the age and with the link that exist with the local environment (e.g. school), this findings appear quite reasonable. As disrupting from school can have a strong impact on the well-being of children, also single households may incur in large psychological costs when they have to leave from their friends, neighbors, family, etc. This could explain why single households, which are generally seen as more flexible in their residential decision, are found to be particularly adverse to relocation.

Furthermore, the study provides a number of interesting findings on how heterogeneous households value locational characteristics. It results for instance that singles and younger households have a relatively strong preference for locations where consumption and recreational amenities are largely available. This supports the idea that such types of households have a higher demand for social contacts and for large marriage markets (see e.g. Glaeser et al., 2001). Couples and elderly households, on the other hand, tend to reside in less "vibrant" locations, where the natural landscape is abundant and the local climate is less congested. Regarding couples, moreover, the co-location problem that is generally ascribed to them seems to play a role also in our study. This is revealed in

particular by the remarkably high MWTP of couples for accessibility to employment and transport infrastructure couples.

We find that also the presence of children has a significant impact on the household preference profile. It results indeed that this category prefers to reside close to places where open space and services are more abundant. Considering the tighter time constraint that expectably involves the parents, this might imply that they have a stronger taste for larger, “multi-functional” locations. For what concerns highly educated households, the study shows that they prefer to live away from the locations where shops and services are largely available. Moreover, it is quite clear from the results that they strongly prefer municipalities providing a good accessibility to employment and to transport facilities. We claim that higher educated households might prefer to reside in municipalities surrounding the main urban centers because, while having a relatively easy access to both consumption and employment areas, such locations allow them to enjoy various comforts that are generally poorly available in big cities (less congestion, houses with garden and parking space, etc). Controlling for the level of amenities available in neighboring municipalities appear advisable for the analysis of phenomena pertaining to small territories such as SNB.

On the base of our findings, the study provides relevant evidences both for scholars and policy makers. Moving costs are found to play a major role in households' residential choices, so overlooking frictions in the housing market could lead to misleading conclusions on agents behaviours and on their preferences for local amenities. On the other hand, the structure of households' MWTP provided in the paper can be interesting for local planners. Having information on the demand of specific groups of dwellers could be useful if one wants to increase the effectiveness of the local policies. For example, promoting the supply of consumption and recreational amenities in main cities would be a good way to attract young households, and make the local social atmosphere more amusing and dynamic. In large cities the investments on education and other public services could be useful not only to attract young households and knowledge workers, but also to accommodate the demand of households with children. Considering the importance that mobility and accessibility has for the highly educated, a viable strategy to attract them would be to improve transport facilities, e.g. by investing in the road infrastructure and in public transport. Given the lack of historical amenities characterizing SNB, and considering the relatively high levels of congestion characterizing most cities in the Randstad, the SNB stakeholders should consider this type of policy with particular regard.

References

- Acemoglu, D., & Angrist, J. (2001). *How large are human-capital externalities? evidence from compulsory-schooling laws*. In *NBER Macroeconomics Annual 2000*, Volume 15 (pp. 9-74). MIT Press.
- Bayer, P., McMillan, R., & Rueben, K. (2004). *An equilibrium model of sorting in an urban housing market* (No. w10865). National Bureau of Economic Research.
- Bayer, P., Ferreira, F., & McMillan, R. (2007). *A unified framework for measuring preferences for schools and neighborhoods* (No. w13236). National Bureau of Economic Research.
- Bayer, P., Keohane, N., & Timmins, C. (2009). *Migration and hedonic valuation: The case of air quality*. *Journal of Environmental Economics and Management*, 58(1), 1-14.
- Bayer, P., McMillan, R., Murphy, A., Timmins, C. (2010), *A Dynamic Model of Demand for Houses and Neighbourhoods*, (No. w17250). National Bureau of Economic Research.
- Bayer, P., & McMillan, R. (2012). *Tiebout sorting and neighborhood stratification*. *Journal of Public Economics*, 96(11), 1129-1143.
- Berry, S. T. (1994). *Estimating discrete-choice models of product differentiation*. *The RAND Journal of Economics*, 242-262.
- Brainport Eindhoven (2012), *Brainport Monitor 2012 – The Power of the Smartest*, <http://www.brainport.nl/en/facts-figures/bp-monitor2012regio-eng-web>.
- Brainport Eindhoven (2005), *Brainport Navigator 2013 – Beyond Lisbon!*, http://archieff.brainport.nl/Brainport_C01/default.asp?comid=32&modid=2099&itemid=0&time=-1.
- Brueckner, J. K., Thisse, J. F., & Zenou, Y. (1999). *Why is central Paris rich and downtown Detroit poor?: An amenity-based theory*. *European Economic Review*, 43(1), 91-107.
- Caetano, G. (2010). *Identification and Estimation of Parental Valuation of School Quality in the US*. *Unpublished Manuscript*.
- Carlino, G. A., & Saiz, A. (2008). *Beautiful city: Leisure amenities and urban growth*. FRB of Philadelphia Working Paper No. 08-22
- Costa, D. L., & Kahn, M. E. (2000). *Power couples: changes in the locational choice of the college educated, 1940–1990*. *The Quarterly Journal of Economics*, 115(4), 1287-1315.

- Duranton, G., & Puga, D. (2001). *Nursery cities: Urban diversity, process innovation, and the life cycle of products*. *American Economic Review*, 1454-1477.
- Ellison, G., Glaeser, E. L., & Kerr, W. (2007). *What causes industry agglomeration? Evidence from coagglomeration patterns* (No. w13068). National Bureau of Economic Research.
- Epple, D. (1987). *Hedonic prices and implicit markets: estimating demand and supply functions for differentiated products*. *The Journal of Political Economy*, 95(1), 59-80.
- Epple, D., Romano, R., & Sieg, H. (2012). *The intergenerational conflict over the provision of public education*. *Journal of Public Economics*, 96(3), 255-268.
- Florida, R. (2004). *The rise of the creative class and how it's transforming work, leisure, community and everyday life*. New York: Basic Books.
- Florida, R., Mellander, C., & Stolarick, K. (2011). *Beautiful places: the role of perceived aesthetic beauty in community satisfaction*. *Regional Studies*, 45(1), 33-48.
- Glaeser, E. L., Scheinkman, J., & Shleifer, A. (1995). *Economic growth in a cross-section of cities*. *Journal of Monetary Economics*, 36(1), 117-143.
- Glaeser, E. L., Kolko, J., & Saiz, A. (2001). *Consumer city*. *Journal of economic geography*, 1(1), 27-50.
- Glaeser, E. L., & Saiz, A. (2003). *The rise of the skilled city* (No. w10191). National Bureau of Economic Research.
- Gottlieb, J. D., & Glaeser, E. L. (2006). *Urban resurgence and the consumer city*. *Urban Studies*, 43(8), 1275-1299.
- Glaeser, E. L., & Gottlieb, J. D. (2009). *The wealth of cities: Agglomeration economies and spatial equilibrium in the United States* (No. w14806). National Bureau of Economic Research.
- Henderson, J. V. (2003). *Marshall's scale economies*. *Journal of urban economics*, 53(1), 1-28.
- Allen Klaiber, H., & Phaneuf, D. J. (2010). *Valuing open space in a residential sorting model of the Twin Cities*. *Journal of Environmental Economics and Management*, 60(2), 57-77.
- Kuminoff, N. V., Smith, V. K., & Timmins, C. (2010). *The new economics of equilibrium sorting and its transformational role for policy evaluation* (No. w16349). National Bureau of Economic Research.

Marshall, A. (1920). *Principles of economics: an introductory volume*. London: MacMillan.

McFadden, D. (1973). *Conditional logit analysis of qualitative choice behaviour* (pp. 105-102). *Frontiers of econometrics*. New York: Academic Press.

McFadden, D. (1978). *Modelling the choice of residential location* (pp. 75-96). Institute of Transportation Studies, University of California.

Moretti, E. (2004). *Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data*. *Journal of econometrics*, 121(1), 175-212.

Moretti, E. (2012). *The new geography of jobs*. New York: Houghton Mifflin Harcourt.

Muth, R. (1969). *Cities and housing: The spatial patterns of urban residential land use*. University of Chicago, Chicago.

Rauch, J. E. (1991). *Productivity gains from geographic concentration of human capital: evidence from the cities* (No. w3905). National Bureau of Economic Research.

Roback, J. (1982). *Wages, rents, and the quality of life*. *The Journal of Political Economy*, 1257-1278.

Rosen, S. (1974). *Hedonic prices and implicit markets: product differentiation in pure competition*. *The journal of political economy*, 82(1), 34-55.

Rouwendal, J., & Thomese, F. (2010). *Homeownership and demand for long-term care*. Tinbergen Institute Discussion Paper TI2010-114/3, Tinbergen.

Rouwendal, J. (2009). *Housing wealth and household portfolios in an ageing society*. *De Economist*, 157(1), 1-48.

Sheppard, S. (1999). *Hedonic analysis of housing markets*. *Handbook of regional and urban economics*, 3, 1595-1635.

Simon, C. J., & Nardinelli, C. (2002). *Human capital and the rise of American cities, 1900–1990*. *Regional Science and Urban Economics*, 32(1), 59-96.

Tiebout, C. M. (1956). *A pure theory of local expenditures*. *The journal of political economy*, 64(5), 416-424.

Van Duijn, M., & Rouwendal, J. (2013). *Cultural heritage and the location choice of Dutch households in a residential sorting model*. *Journal of Economic Geography*, 13(3), 473-500.

Appendix A. Supplement tables

Table A.1. First stage results for homeowners: estimates for interaction parameters and mean effects moving costs with single coefficient

Municipality characteristics	Household characteristics				
	Mean	Highly educated households	Age (+10 years)	Households with children (-18)	Couples
Standardized house price (in euros)	(1)	(1)	(1)	(1)	(1)
Nr. of shops		10.25 (-0.6)***	-0.0493(-0.0172)***	4.576 (-0.545)***	-5.784 (-0.642)***
Nature (km ²)		-0.007 (-0.000174)***	-0.000232 (-0.0000053)***	0.000373 (-0.000166)**	-0.00615 (-0.00019)***
Water(km ²)		-0.0452 (-0.00212)***	-0.00114 (-0.0000581)***	-0.0147 (-0.00193)***	-0.0166 (-0.00222)***
Distance to 100,000 jobs (km)		-0.519 (-0.0283)***	-0.0157 (-0.000831)***	0.276 (-0.0262)***	-0.019 (-0.0319)***
Distance to intercity station (km)		0.0291 (-0.0141)**	0.000831 (-0.00041)***	0.162 (-0.013)***	-0.352 (-0.0154)***
Distance to motorway ramp		-0.74 (-0.0125)***	-0.000616(-0.000349)*	-0.0855 (-0.0109)***	0.259 (-0.0134)***
		-0.0516 (-0.0078)***	-0.00773 (-0.000227)***	-0.0085 (-0.00703)	0.0864 (-0.00842)***
Moving costs homeowner		-7.964 (-0.0189)***			
Moving costs renter		-6.304 (-0.0165)***			

The table reports the estimates for the interaction parameters and for the single parameter referring to moving costs. The moving cost parameters indicates the estimated average impact on the indirect utility respectively for homeowners and for renters. Variables are normalized to have mean zero, so the coefficients represent the estimated deviation from the mean indirect utility. Standard errors in parenthesis. 90, 95 and 99% significance level are indicated, respectively, as *, ** and ***. Dummies for Eindhoven and Helmond are included in the estimation. Movers are identified by households who moved to their 2012 location later

Table A.2. Estimates of mean indirect utilities corresponding to different model specifications

Municipality	Scenario 1 (1)	Scenario 2 (2)	Scenario 1 (3)
Asten	-1.598 (-0.0172)***	-1.835 (-0.047)***	-2.158 (-0.052)***
Best	-0.867 (-0.0135)***	-1.247 (-0.0392)***	-1.694 (-0.0425)***
Deurne	-0.985 (-0.0141)***	-0.153 (-0.0348)***	-0.284 (-0.0373)***
Eersel	-1.231 (-0.0151)***	-1.057 (-0.0417)***	-1.352 (-0.0444)***
Eindhoven	0.776 (-0.00912)***	1.331 (-0.0253)***	1.200 (-0.0269)***
Nuenen, Gerwen en Nederwetten	-1.171 (-0.0145)***	-1.305 (-0.0412)***	-1.362 (-0.0434)***
Oirschot	-1.326 (-0.0158)***	-0.727 (-0.0416)***	-0.937 (-0.0453)***
Someren	-1.450 (-0.0162)***	-0.935 (-0.0414)***	-1.139 (-0.043)***
Son en Breugel	-1.482 (-0.0164)***	-0.603 (-0.0389)***	-0.853 (-0.0413)***
Valkenswaard	-0.910 (-0.014)***	-0.157 (-0.0341)***	-0.784 (-0.041)***
Veldhoven	-0.550 (-0.0123)***	-0.634 (-0.0355)***	-0.877 (-0.0378)***
Waarle	-1.431 (-0.0163)***	-1.281 (-0.0451)***	-1.693 (-0.05)***
Gemert-Bakel	-0.885 (-0.014)***	-0.179 (-0.0364)***	-0.563 (-0.0416)***
Heeze-Leende	-1.542 (-0.0175)***	-0.203 (-0.0375)***	-0.541 (-0.0401)***
Laarbek	-1.161 (-0.0148)***	-0.845 (-0.04)***	-1.251 (-0.0439)***
Reusel-De Mierden	-2.056 (-0.0206)***	-1.490 (-0.0447)***	-2.030 (-0.0507)***
Cranendonk	-1.240 (-0.0157)***	-0.549 (-0.0376)***	-0.958 (-0.0425)***
Bergeijk	-1.379 (-0.016)***	-0.604 (-0.0384)***	-0.795 (-0.0407)***
Bladel	-1.348 (-0.016)***	-1.217 (-0.0418)***	-1.369 (-0.0437)***
Geldrop-Mierlo	-0.643 (-0.0126)***	-0.419 (-0.0355)***	-0.761 (-0.0376)***

Reference municipality is Helmond. The table reports the sets of relative mean indirect utilities estimated for the owner-occupied market under three different scenarios. Respectively: scenario 1 refers to the parameters estimated in the first step regression of the residential sorting model when moving costs are not included (i.e. the estimates reported in column 1 of Table 6); scenario 2 refers to the parameters estimated in the first step regression when moving costs are included with only a single coefficient (i.e. the estimates reported in column 1 of Table A.1.; scenario 3 refers to the parameters estimated in the first step regression when moving costs are included with both an interaction and single coefficient (i.e. the estimates reported in column 2 of Table 6).

Table A.3.a. Correlation matrix mean indirect utilities

	Scenario 1	Scenario 2	Scenario 3
Scenario 1	1		
Scenario 2	0.770	1	
Scenario 3	0.779	0.980	1

Table A.3.b. Correlation matrix interaction parameters

	Scenario 1	Scenario 2	Scenario 3
Scenario 1	1		
Scenario 2	0.034	1	
Scenario 3	0.032	0.989	1

Table A.3.c. Correlation matrix all parameters

	Scenario 1	Scenario 2	Scenario 3
Scenario 1	1		
Scenario 2	0.000	1	
Scenario 3	-0.012	0.986	1

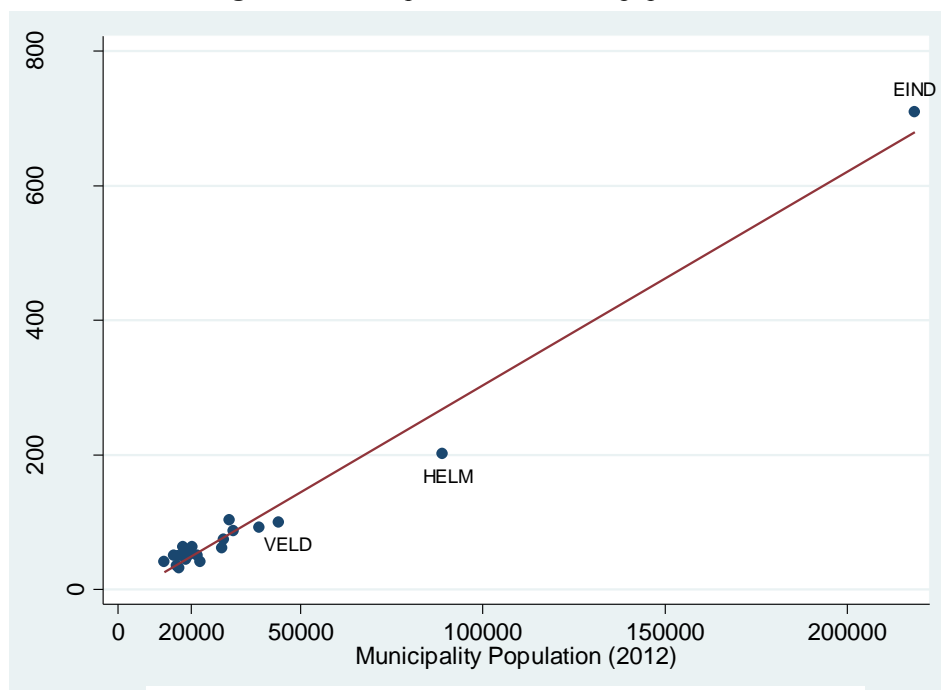
The three tables report the correlation coefficients obtained by comparing sets of parameters estimated under three different scenarios. Respectively: scenario 1 refers to the first step estimates with no moving costs (i.e. the estimates reported in column 1 of Table 6); scenario 2 refers to the first step estimates with moving costs included with single coefficient (i.e. the estimates reported in column 1 of Table A.1.); scenario 3 refers to the first step estimates with moving costs included with both an interaction and single coefficient (i.e. the estimates reported in column 2 of Table 6). All correlation coefficients are estimated considering the parameter of the homeowners only. Tab A.3.a reports the correlation among the vectors of mean indirect utilities reported in Table A.2. Table 5b reports the correlation among the interaction parameters derived in the first stage estimation. Table A.3.c reports the correlation coefficient estimated by considering both the mean indirect utilities and the interaction parameters.

Table A.4. Second stage estimates

VARIABLES	(1) OLS	(2) 2SLS	(3) 2SLS
Standardized house price (in euros)	-0.673 (-0.181)***	-14.95 (-3.524)***	-14.91 (-3.508)***
Nr. of shops	0.00133 (-0.00043)***	0.000935 (-0.000614)	0.000969 (-0.000636)
Distance to 100,000 jobs (km)	-0.0246 (-0.00743)***	-0.291 (-0.077)***	-0.290 (-0.0763)***
Distance to intercity station (km)	-0.0257 (-0.00553)***	-0.025 (-0.0294)	-0.0249 (-0.0295)
Distance to motorway ramp	0.00048 (-0.00709)	-0.0657 (-0.0382)*	-0.0656 (-0.0381)*
Nature (km²)	0.00507 (-0.00218)**	0.0625 (-0.0178)***	0.0624 (-0.0177)***
Water(km²)	0.00252 (-0.00183)	0.0278 (-0.00988)***	0.0277 (-0.00985)***
Constant	9.597 (-2.263)***	186.3 (-43.65)***	185.8 (-43.45)***
IV housing price	no	yes	Yes
IV shops	no	no	Yes

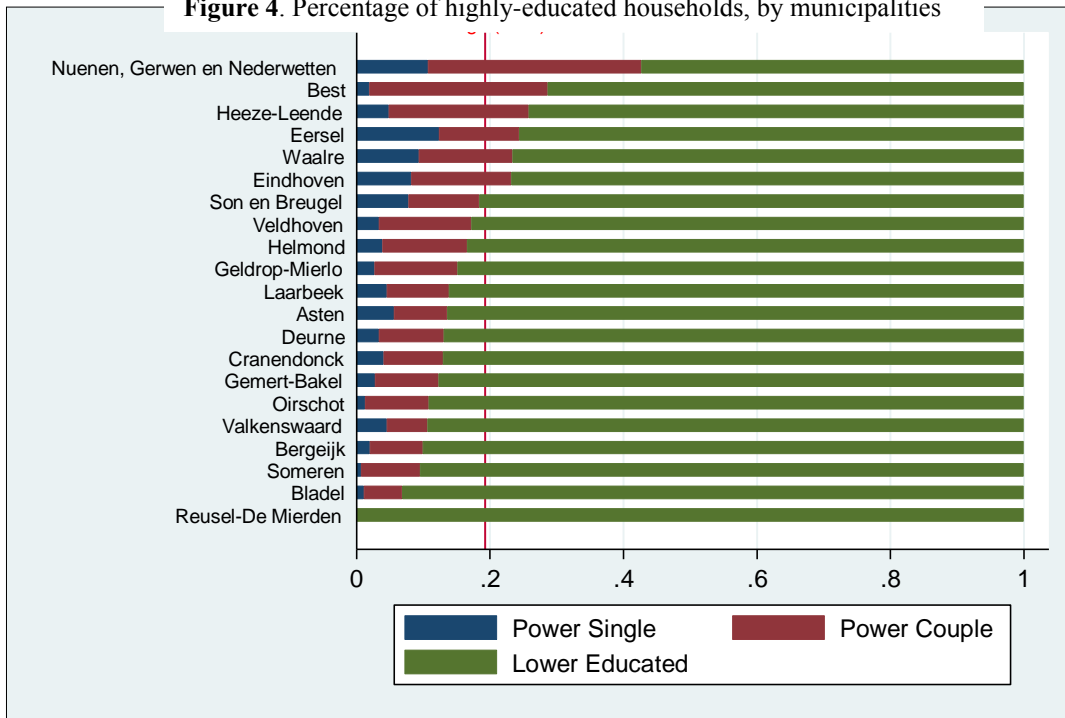
Estimates from 399 observations. Standard errors in parenthesis. 90, 95 and 99% significance level are indicated, respectively, as *, ** and ***.

Appendix B. Supplement figures

Figure 3. Consumption amenities and population

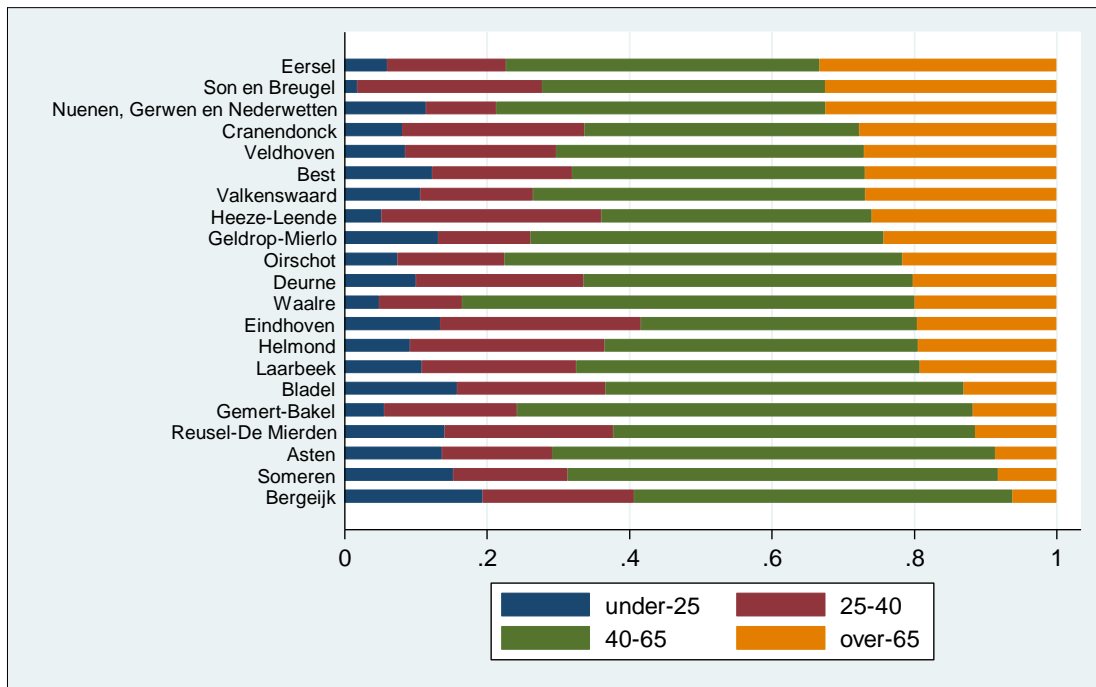
Source: ABF(2007), CBS(2012)

Figure 4. Percentage of highly-educated households, by municipalities



Source: WoON 2012

Figure 5. Demographic composition (age), by municipality



Source: WoON 2012

Appendix C. Derivation of marginal willingness to pay

In order to derive the MWTP, it is useful to consider the extended form of the indirect utility function specified in equation 4, and distinguish the housing prices from the other municipality characteristics:

$$V_{i,n} = \beta_{0,p} p_{k,n} + \beta_{p,l} (Z_{i,l} - \bar{Z}_l) p_{k,n} + \sum_{k=1}^K \beta_{0,k} X_{k,n} + \xi_n + \sum_{k=1}^K \left(\sum_{l=1}^L \beta_{k,l} (Z_{i,l} - \bar{Z}_l) \right) X_{k,n} - \alpha_i^{MC} MC * I_{i,n} + \varepsilon_{i,n}$$

Then we take to the left-hand side the price variable $p_{k,n}$, where $p_{k,n} = \ln P_{k,n}$, and $V_{i,n}$ to the right hand side:

$$-\left(\beta_{0,p} + \beta_{p,l} (Z_{i,l} - \bar{Z}_l) \right) p_{k,n} = \sum_{k=1}^K \beta_{0,k} X_{k,n} + \sum_{k=1}^K \left(\sum_{l=1}^L \beta_{k,l} (Z_{i,l} - \bar{Z}_l) \right) X_{k,n} + \varepsilon_{i,n} + \xi_n - \left(\gamma_0 + \sum_{l=1}^L \gamma_l (Z_{i,l} - \bar{Z}_l) \right) MC * I_{i,n} - V_{i,n}$$

Rearranging, we can rewrite this expression as follows:

$$p_{k,n} = - \left(\frac{\beta_{0,k} + \beta_{k,l} (Z_{i,l} - \bar{Z}_l)}{\beta_{0,p} + \beta_{p,l} (Z_{i,l} - \bar{Z}_l)} \right) X_{k,n} - \left(\frac{1}{\beta_{0,p} + \beta_{p,l} (Z_{i,l} - \bar{Z}_l)} \right) \varepsilon_{i,n} - \left(\frac{1}{\beta_{0,p} + \beta_{p,l} (Z_{i,l} - \bar{Z}_l)} \right) \xi_n + \left(\frac{1}{\beta_{0,p} + \beta_{p,l} (Z_{i,l} - \bar{Z}_l)} \right) V_{i,n} + \left(\frac{1}{\gamma_0 + \sum_{l=1}^L \gamma_l (Z_{i,l} - \bar{Z}_l)} \right) MC * I_{i,n}$$

Hence, given that $p_{k,n} = \ln P_{k,n}$ the MWTP for locational characteristics can be derived by taking the partial derivative of the function above:

$$\frac{\partial P_{k,n}}{\partial X_{k,n}} = - \left(\frac{\beta_{0,k} + \beta_{k,l} (Z_{i,l} - \bar{Z}_l)}{\beta_{0,p} + \beta_{p,l} (Z_{i,l} - \bar{Z}_l)} \right) P_{k,n}$$

As households characteristics are normalized so as to have mean zero, the MWTP for the average households can be expressed simply as:

$$\frac{\partial P_{k,n}}{\partial X_{k,n}} = - \left(\frac{\beta_{0,k}}{\beta_{0,p}} \right) P_{k,n}$$

Similarly than we did with locational characteristics, the MWTP referring to the moving cost component will be given by:

$$\frac{\partial P_{k,n}}{\partial I_{i,n}} = - \left(\frac{\gamma_0 + \sum_{l=1}^L \gamma_l (Z_{i,l} - \bar{Z}_l)}{\beta_{0,p} + \beta_{p,l} (Z_{i,l} - \bar{Z}_l)} \right) P_{k,n} ; \quad \frac{\partial P_{k,n}}{\partial I_{i,n}} = - \left(\frac{\gamma_0}{\beta_{0,p}} \right) P_{k,n}$$