

van Duijn, Mark; Rouwendal, Jan

**Conference Paper**

## Sorting based on amenities and income composition: Evidence on the multiplier effect

53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* van Duijn, Mark; Rouwendal, Jan (2013) : Sorting based on amenities and income composition: Evidence on the multiplier effect, 53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/123990>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Sorting based on amenities and income composition: Evidence on the multiplier effect

Mark van Duijn<sup>\*,†</sup> & Jan Rouwendal<sup>\*,\*\*</sup>

\* *Department of Spatial Economics, VU University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands*

\*\* *Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, The Netherlands*

† *Corresponding author: [m.van.duijn@vu.nl](mailto:m.van.duijn@vu.nl)*

29 June 2013

**Abstract.** Recent research shows that urban heritage can be an important factor of residential location choice, and that especially highly educated households are attracted by the presence of cultural heritage. In this paper, we use a location choice model to investigate the attractiveness of neighborhood amenities for homeowners with a higher income in Amsterdam and surrounding municipalities. We control for observed (house prices, proximity to large labor markets, area of historic city center, share of rich households) and unobserved neighborhood characteristics. Since the attractiveness of neighborhoods may be affected by amenities of surrounding neighborhoods, we extend the model by incorporating these spillover effects. We use a unique data set where we identify the marginal willingness-to-pay of these different households for living within or near the historic city center. The results show that high income households not only prefer to live close to each other, but also prefer to live close to the historic city center. Also the number of shops seem to be higher in the historic city center. This suggests that there are multiplier effects regarding the effect of historic city centers.

**Key words:** Location choice, cultural heritage, multiplier effect, equilibrium sorting models, discrete choice models, unobserved characteristics, heterogeneous household preferences

**JEL classifications:** R2, J1, Z1

## 1 Introduction

Urban amenities are becoming more prominent in the residential location choice. Although the work location will always be an important factor for location choice of households, as is stressed in the Alonso-Muth-Mills models, other consumer needs are growing in relative importance. This has certain implications for urban growth. It is argued that certain consumers enjoy urban heritage, and therefore they choose a location that has a cultural identity. If this statement is correct, research on which consumers are attracted to these locations is very interesting for local policy makers. This is not only relevant for the largest European cities, of which most have a historic city center, but also for other cities that have the prospect of becoming a historic city.

Typical consumers can be identified, for example, by their current income, life-cycle status, and whether they own or rent their home. This paper investigates the residential preferences of these types of households by analyzing their location choice. We investigate the heterogeneous preferences of these households on neighborhood characteristics, such as house prices, proximity to large labor markets, concentration of high income households, and the area of historic city center. We focus on the Amsterdam area. We use a horizontal sorting model, following Bayer, McMillan & Rueben (2004; 2007), and Bayer & Timmins (2005; 2007), to find evidence on which amenities drive sorting considering different types of households. The basic idea of the sorting model is that households choose among neighborhoods on the basis of the neighborhood characteristics. We do not observe all observed characteristics. The advantage of using a sorting model is that it controls for heterogeneous households and unobserved characteristics.

Urban heritage is one of the observed amenities. We define it as all those features (e.g. listed buildings, monuments) that relate the past to the present, but it is the combination of all these features which contribute to the cultural identity of a city. In the literature, this is called the *ensemble* effect (Brueckner, Thisse & Zenou, 1999; Lazrak et al., 2011). In this paper we consider the historic city center of Amsterdam, which is a national conservation area, where most important historic buildings date back to the seventeenth century – the Dutch Golden Age – or earlier.

This paper also considers including spatial elements in the sorting model by accounting for the characteristics of surrounding neighborhoods. Households living just outside the historic city center can still enjoy the amenities that the historic city center has to offer. This means that certain characteristics of surrounding neighborhoods have an effect on the attractiveness of its own neighborhood. Taking into account the spatial elements in the sorting model can therefore be of potential importance. Hence our modeling approach combines an equilibrium sorting model and spatial spillover effects.

This paper is related to Duijn & Rouwendal (2013) who study the distribution of heterogeneous households over Dutch municipalities using a sorting model. In this paper, we concentrate on a smaller spatial scale, viz. the Amsterdam area. Since this area can be regarded as a single labor market, we can take wages as given. Since the spatial units that we now distinguish are smaller, we pay more attention to the impact of the demographic composition on neighborhood choice. We focus, in particular, on the concentration of high income households. We believe there is a relation between the historic city center, its shops and the concentration of high income households in the Amsterdam area. We provide strong evidence that high income homeowners have a higher willingness-to-pay for living in and around the historic city center and in neighborhoods with a high percentage of other high income households. If these high income homeowners indeed choose to live in or a close to the historic city center, this increases the share of high income households in the neighborhood, which then attracts additional high income households. This suggests that there are multiplier effects present regarding the effect of living in or close to a historic city center. We test in a simple way whether this is the case.

In Section 2, we discuss the methodology concerning the residential sorting model and the introduction of spatial elements in the model. Our unique data on household and neighborhood characteristics, as well as, some descriptive statistics are discussed in Section 3. Section 4 is devoted to the estimation results. The implications of these results are discussed in Section 5. The multiplier effect is researched in Section 6. Section 7 summarizes and concludes.

## **2 The location choice model**

### **2.1 Methodology**

This study focuses on the location decisions of households with a different income composition and whether these households are attracted by cultural heritage within the Amsterdam area. We observe large variation in house prices within the Amsterdam area (see Figure A.1 in Appendix A). If this picture also reflects the income distribution, it is interesting to investigate these differences. We distinguish households by their income, their social economic status and whether they own or rent a house. Apart from the quality of the house, it is interesting to analyze what other motives play a role in the location decision of these households. Although, we are in particular interested which of these households are attracted by cultural heritage, we also investigate if they are attracted by neighborhoods with a high share of high income households.

We focus on the historic city center of Amsterdam as an amenity that plays a role in the location choice of households. The historic city center contributes to the atmosphere in the area and to its

attractiveness for residents, firms, and tourists. As a result, there are more shops, restaurants and similar endogenous amenities in these areas, which further contribute to its attractiveness. Hence, it may be the case that historic city centers have a multiplier effect through its impact on endogenous amenities. We provide a first step to identify those multiplier effects regarding the effect of the historic city center with respect to the concentration of high income households (for technical details, see Appendix B). We do this by studying the outcomes of a structural model, the sorting model. The multiplier effect is then further analyzed by regressing the number of shops in each neighborhood to the important location choice factors that are identified by the sorting model.

We use the framework of sorting models developed by Bayer, McMillan & Rueben (2004). This framework has emerged in the last two decades and has its roots in the theoretical literature (e.g. Tiebout, 1956; Epple, Filimon & Romer, 1984; Benabou, 1993; Anas & Kim, 1996; Nechyba, 1999) analyzing how households sort themselves into local jurisdictions to enjoy its desired level of a public good. In the empirical literature, two main types of household location choice models – horizontal and vertical – are distinguished that diverge in the type of heterogeneity in preferences they allow. Vertical sorting models often study heterogeneity in a single dimension (e.g. Epple & Platt, 1998; Epple & Sieg, 1999). Horizontal sorting models use the additive random utility framework for discrete choice, first introduced by McFadden (1973), that allows for a more flexible approach concerning amenity and household heterogeneity (which is particularly relevant in our case). Bayer, McMillan & Rueben (2004) provide a framework for applying the horizontal sorting model which has recently been applied in a variety of empirical studies (Timmins, 2005; Murdock, 2006; Klaiber & Phaneuf, 2010; Van Duijn & Rouwendal, 2013). For applications with a large number of choice alternatives, the multinomial logit (MNL) is the only tractable specification of an additive random utility model. Although this model is known to suffer from the restrictive independence of irrelevant alternatives (IIA) property. Recent literature has stressed that this property is not present at the level of the population if a sufficient amount of heterogeneity among the actors is allowed.<sup>1</sup> Moreover, it is now possible to incorporate unobserved characteristics of choice alternatives into the model, and to deal with the related endogeneity problems through a two step procedure.

---

<sup>1</sup> For an insightful discussion on the assumptions of the vertical and horizontal sorting models, see Kuminoff, Smith & Timmins (2010) and Van Duijn & Rouwendal (2013). Another way to deal with heterogeneity among actors are the mixed logit models (see, for example, Train & McFadden, 2000), which we will not discuss here.

## 2.2 Basic specification

The framework that we use in this paper follows the sorting model developed by Bayer, McMillan & Rueben (2004). We follow the random utility framework and consider a population of households, indexed by  $i = 1 \dots I$ , that have to choose a residential location from a large number,  $N$ , of alternatives  $n = 1 \dots N$  with  $K$   $k = 1 \dots K$  characteristics. In our application, the alternatives consist of neighborhoods in the Amsterdam area. The utility of household  $i$  from choosing to reside in neighborhood  $n$  is given as:

$$u_{i,n} = \sum_{k=1}^K \alpha_{i,k} X_{k,n} + \xi_n + \varepsilon_{i,n}. \quad (2.1)$$

In this equation,  $X_{k,n}$  denotes the value of the  $k$ -th characteristic of neighborhood  $n$ , the vector  $\xi_n$  are the unobserved neighborhood characteristics, the  $\alpha_{i,k}$ 's are coefficients and the  $\varepsilon$ 's random variables. The vector  $X_n$  includes indicators of observable neighborhood characteristics, observable household characteristics, and prices. Note that the coefficients,  $\alpha_{i,k}$ , are individual-specific. They are functions of household characteristics  $Z$ :

$$\alpha_{i,k} = \beta_{0,k} + \sum_{l=1}^L \beta_{k,l} Z_{i,l} - Z_l. \quad (2.2)$$

In this equation  $Z_{i,l}$  denotes the value of the  $l$ -th characteristic of household  $i$   $l = 1 \dots L$ , and  $Z_l$  the population mean of characteristic  $l$ , while the  $\beta$ 's are coefficients. Household characteristics include income, life-cycle status, et cetera. Using Equation 2.2, we can rewrite utility as:

$$u_{i,n} = \sum_{k=1}^K \beta_{0,k} X_{k,n} + \xi_n + \sum_{k=1}^K \sum_{l=1}^L \beta_{k,l} Z_{i,l} - Z_l X_{k,n} + \varepsilon_{i,n}. \quad (2.3)$$

If we assume that all random variables are IID extreme value type I distributed, the choice probabilities of utility maximizing households can be derived in closed form as:

$$Pr_{i,n} = \frac{\exp w_{i,n}}{\sum_{m=1}^M \exp w_{i,m}}, \quad (2.4)$$

where  $w_n^i$  is the deterministic part of the utility function in Equation 2.3. The sum of the probabilities yields the predicted demand for each neighborhood. Given the housing stock in each neighborhood, demand should be equal to the supply in each neighborhood to clear the housing market. The housing market clearing condition is then defined as:

$$\sum_{i=1}^I Pr_{i,n} = S_n. \quad (2.5)$$

It is important to note that this condition determines the equilibrium price of housing in each neighborhood. Although the model does not result in a closed-form specification of the equilibrium price equation, estimated versions allow the computation of equilibrium prices for all neighborhoods with counterfactual values of amenities and/or housing supply in some or all neighborhoods. Moreover, substitution of these equilibrium prices into the choice probability equation  $Pr_{i,n}$  allows one to study the change in the demographic composition of the neighborhood population induced by the change in amenities. This will be discussed in detail later on.

An important property of the model is that it accounts for unobserved heterogeneity in neighborhoods,  $\xi_n$ . In practice, a researcher is incompletely informed about the characteristics of a neighborhood that are relevant for household welfare. Ignoring the unobserved heterogeneity in neighborhoods may, in practice, have modest consequences as long as  $\xi$  is uncorrelated with the  $X$ 's. Since the housing price is one of the  $X$ 's, we have good reasons to think that unobserved neighborhood characteristics have an effect on house prices. It may be the case that some households choose a particular neighborhood where housing is expensive because of the presence of an attractive (unobserved) amenity that makes it well worth paying the higher price. Similarly, it may be the case that some households are reluctant to choose a neighborhood with a low housing price because of a negative (unobserved) amenity. This problem was addressed rigorously by Berry (1994) and Berry, Levingsohn & Pakes (1995) in the context of the automobile market. Their solution is to estimate the model (Equation 2.3) in a two-step estimation procedure. To explain their method, we start by rewriting Equation 2.3 as:

$$u_{i,n} = \delta_n + \sum_{k=1}^K \sum_{l=1}^L \beta_{k,l} Z_{i,l} - Z_l X_{k,n} + \varepsilon_{i,n} \quad (2.6)$$

with

$$\delta_n = \sum_{k=1}^K \beta_{0,k} X_{k,n} + \xi_n \quad (2.7)$$

In the first step one estimates the vector of mean indirect utilities,  $\delta$ , and the coefficients of the cross effects of household and alternative characteristics,  $\beta_{k,l}$ , on the basis of Equation 2.6. These are estimated as a MNL, in which the  $\delta$ 's are specified as alternative specific constants. The MNL model predicts the probability,  $Pr_{i,n}$ , that each household  $i$  chooses alternative  $n$ , as is illustrated in Equation 2.4. The sum of the probabilities for each alternative are forced to be equal to the housing stock for each alternative by adjusting the alternative specific constants – hence satisfying the equilibrium constraint (Equation 2.5).

In the second step the  $\delta_n$ 's are analyzed further on the basis of Equation 2.7. Because of the endogeneity problem just discussed, using OLS provides biased coefficients of the  $\beta_{0,k}$ 's.<sup>2</sup> However, 2SLS can be used if an instrument for the price is available. Berry, Levinsohn & Pakes (1995) suggested to use the characteristics of similar alternatives as instruments, but in one of the first applications of this methodology to sorting on the housing market Bayer, McMillan & Rueben (2004) proposed a different approach, which we follow in our analysis. We construct a single instrument by solving for the vector of prices that would clear the market if there were no unobserved heterogeneity (that is if all  $\xi_n$ 's were equal to zero). This instrument is by construction independent of the unobserved heterogeneity term  $\xi$  and in all probability strongly correlated with the observed housing prices.

Some additional explanation on the creation of the instrument is useful. The instrument should be based on the 'true' values of the coefficients in Equation 2.3. However, initially the  $\beta_{0,k}$ 's are still unknown and we, therefore use an iterative procedure. Initial values of the  $\beta_{0,k}$ 's are obtained by estimating Equation 2.7 via OLS. These coefficients, along with those obtained from estimating the MNL (Equation 2.6), are then used to calculate a price vector,  $p$ , that, after setting  $\xi_n = 0$  for all  $n$ , satisfies the equilibrium condition (Equation 2.5). This price vector is then used as an instrument for prices by estimating Equation 2.7 via 2SLS. This results in new coefficients for  $\beta_{0,k}$  which we plug back into Equation 2.3 to solve for a new price vector that satisfies the equilibrium condition (Equation 2.5) in the same way as before. This process is repeated until the instrument stabilizes.<sup>3</sup>

We also include a neighborhood socio-demographic characteristic in the utility function (percentage of rich households), which is likely to be endogenous with respect to the unobserved neighborhood characteristics. We will address this issue in our empirical work. Earlier research, for instance Bayer, Ferreira & McMillan (2007), has shown that households have strong preferences for living with similar households that are similar in income, ethnicity or education. To take this into account we introduce aspects of the demographic composition of neighborhoods that are potentially relevant into the household utility function that we estimate. The demographic composition is therefore as much a determinant of the choice behavior as its outcome. To take the associated heterogeneity into account we will use an instrumental variable strategy that is similar to the one just explained for the price and simultaneously estimated. That is, we compute the

---

<sup>2</sup> Note that this endogeneity issue arises only in the second step estimation procedure at the aggregate (neighborhood) level. Individual actors take prices and amenities (observable as well as unobservable) as given. This means that we can estimate the first step using a MNL model, which has the unobservable characteristics included in the mean utility, without instrumenting the price.

<sup>3</sup> The instrument stabilizes rapidly (within five iterations) and is independent on the initial coefficients of  $\beta_{0,k}$ . For more discussion on the instrumental variables strategy, see Bayer et al. (2004).



distribution of the households over the neighborhoods that would be observed if no unobserved heterogeneity were present in the model and use this counterfactual distribution as our instrument for the actual distribution, while computing the price instrument. We are able to do so because our data includes all households from the Amsterdam area. This implies that we estimate  $Pr_{i,n}$  for each household, and therefore we can compute the counterfactual concentration of high income households per neighborhood. This instrument is also by construction independent of the unobserved heterogeneity term  $\xi$  and in all probability strongly correlated with the observed concentration of high income households.

### 2.3 Spatial spillovers

The focus of this paper is on attractiveness of the historic city center on the location choice of households with different income composition. This amenity probably also extend over neighborhood boundaries. The historical city center is not only important for households who are located in this conservation area, but often also for households living in the proximity who like to visit such a center for shopping, dining and recreational purposes. Casual evidence suggests that many people appreciate to live in the proximity of historical city centers so that it can easily be visited, but do not necessarily want to reside in neighborhoods in which it is located, for instance because the neighborhood is too noisy or the available houses are too expensive. Choosing a location in a nearby neighborhood may be a strategy that offers an optimal location choice. These considerations suggest that we should take into account the possibility that the attractiveness of a neighborhood as a residential location is determined in part by the amenities in the surrounding neighborhoods.

To take into account the possibility that the attractiveness of a particular neighborhood is partially determined by the amenities in surrounding neighborhoods, we will extend our basic specification of the sorting model in which not only a measure of neighborhood characteristics of the ‘own’ neighborhood in our model is included, but also by incorporating a weighted sum of neighborhood characteristics in the proximity.<sup>4</sup> More specifically, we use the ‘potential’ formulation:

$$Xsur_n = \sum_{m \in C_n} e^{-\varphi d_{mn}} X_m. \quad (2.8)$$

---

<sup>4</sup> Note that our use of characteristics of surrounding neighborhoods as arguments of the mean utilities invalidates their use as instruments for the price. This underlines the importance of our use of computed instruments suggested by Bayer, McMillan & Rueben (2004).

The variable  $Xsur_n$  is a weighted sum of the measure of neighborhood characteristics  $X_m$  in neighborhoods  $m$  in a set  $C_n$  of neighborhoods surrounding  $n$ , where the weights are defined as an exponential function of the distance  $d_{mn}$  between  $m$  and  $n$ . We include both  $Xsur_n$  and  $X_n$  in Equation 2.7. The variable  $Xsur_n$  can also be interpreted as a spatial lag with exponential weights. It introduces a spatial element into the model. Since it relates only to exogenous variables, this has no significant consequences for estimating the model in itself.<sup>5</sup>

### 3 Data and descriptive statistics

We carry out a regional analysis for the Amsterdam area using neighborhoods as our spatial units. The estimation of the sorting model requires two types of data: household and neighborhood characteristics. The location of the household is essential information to combine the household and neighborhood data. We make use of a unique dataset of household characteristics that include all household observations within each neighborhood. In this section, we describe the historic city center, household income and the housing market in the Amsterdam area. Furthermore, we will discuss the data used in the analysis and show some additional descriptive statistics.

#### 3.1 Historic city centers

In the Netherlands, historic city centers are national conservation areas which contain a high concentration of listed buildings and monuments. These areas are designated by the national government for its architectural and historic value. Becoming a conservation area involves a long bureaucratic process that involves many institutions, such as the municipality involved and the Netherlands Institute for Cultural Heritage.<sup>6</sup> In the US, these conservation areas are called historic districts. These are listed on the National Register of Historic Places under the authority of the National Park Service.<sup>7</sup> It is important to note that from the perspective of home owners, the designation of conservation areas in the Netherlands is exogenously determined. Also, the boundaries of the conservation areas do not correspond to the neighborhood boundaries. We can take the historic city center of Amsterdam as an example. Its historic city center is 679 hectare (6.79 km<sup>2</sup>), where the average in the Netherlands is around 75 hectare (0.75 km<sup>2</sup>). Ten

---

<sup>5</sup> The distance decay coefficient,  $\phi$ , is set at 0.5. The function is therefore exponential decreasing and weights are going towards zero when distance increases (weight < 0.1 if distance is 5 km).

<sup>6</sup> 'Rijksdienst voor het Cultureel Erfgoed' in Dutch and this Service is part of the Ministry of Education, Culture and Science of the Netherlands.

<sup>7</sup> The National Park Service is a government office of the United States Department of the Interior. Note that the criteria of designation to become a conservation area could differ between countries.

neighborhoods are within the historic city center of Amsterdam. These neighborhoods are well-known for their canals, gabled houses and numerous monuments. In 2010, the canal ring area inside the Singelgracht (which covers a large part of the historic city center) was added to the UNESCO World Heritage list.

The area of historic city center in each neighborhood and a separate variable for the area of historic city center in surrounding neighborhoods are the indicators for cultural heritage that are included in our analysis. In this paper, the historic city center is measured as the number of square kilometers. We have this information on the neighborhood level. The neighborhood which has the largest part of the Amsterdam historic city center is *Nieuwmarkt en Lastage* (1.03 km<sup>2</sup>). We also include the weighted sum of historic city center in the proximity to estimate the spatial spillover effect, as we discussed in Section 2.3.

### **3.2 Income and the housing market**

The variation in the (gross and net) income in the Netherlands is relatively small compared to many other countries. Within the Netherlands, the larger cities show a higher dispersion between rich and poor than the smaller villages. This is because the Netherlands has a unique rental sector. It has a high share of social housing, in particular in the larger cities. In 2003, around 50% of the housing supply in the city of Amsterdam – compared to the average of 35% in the Netherlands – belonged to the social renting sector. It is therefore not surprising that Amsterdam has a lot of households with a lower income. On the other hand, Amsterdam also has residents with a very high income. If these high income households live in the most expensive houses in Amsterdam, Figure A.1 in Appendix A could give a good picture of where those households reside. Because the owner-occupied and rental sector are very different from each other, we make a distinction between the owner-occupied and rental sector in our analysis. In this way, we should be able to distinguish the preferences of renters and owner-occupiers. We focus on the owner-occupied market because in the rental sector not all households freely choose their location.

### **3.3 Data**

In our application, estimation of the location choice model makes use of household and neighborhood characteristics. Statistics Netherlands provides us with a unique dataset that includes detailed information on all Dutch individuals and households. The information we use in our analysis is from 2010 and contains approximately 600,000 households spread over 320 neighborhoods. Using the information on household characteristics makes it possible to investigate

the heterogeneous preferences of different types of households. We consider gross primary household income, whether households are homeowners or renters, and the social economic status of each household (students, (self-)employed, unemployed and retired households). These information is a snapshot of the status of Dutch households in 2008.

The existing (predominantly Anglo-Saxon) literature mainly focuses on the provision of 'good' schools which is an important determinant of household location choices in the United States (Bénabou, 1996; Fernandez & Rogerson, 1996, 2003; Nechyba, 1999, 2000; Bayer, Ferreira & McMillan, 2004). In the Netherlands, the educational system is different from that in the US and the UK. There are no school districts (households can freely choose a school for their children) and denominational schools are more important than public schools.

The choice set for each household is the Amsterdam area, which includes all neighborhoods in the Amsterdam area. We drop some neighborhoods from the choice set which have almost no household observations, for example neighborhoods with a large share of industrial estate. We have information on the location of each household on the neighborhood level. This enables us to link the household and neighborhood characteristics. We include several neighborhood characteristics in the utility function. We already discussed the historic city center which is likely to be an important amenity for the attractiveness of a neighborhood. We also include the distance to the nearest concentration of 100,000 jobs. This measure is used to reflect the accessibility to jobs. These data are provided by the Netherlands Environment Assessment Agency.<sup>8</sup> The variable reflects the location of the largest agglomeration economies in the Netherlands. In the last decade, there have not been major shifts of large labor markets in the Netherlands. The distance to a high concentration of jobs for the Amsterdam area is low, compared to average the Netherlands. The socio-demographic characteristic of a neighborhood that is included in the analysis is the percentage of high-income households. There is likely correlation between the concentration of high income households and the unobserved neighborhood quality, which causes an endogeneity issue. We proposed an instrumental variable strategy at the end of Section 2.2. We capture the housing market by including a neighborhood price index for a standard house, which we interpret as the price of housing services. The price index is based on estimation of a standard hedonic price method with neighborhood fixed effects. This price is also likely to be correlated with the unobserved neighborhood quality. The instrumental variable strategy we use to account for

---

<sup>8</sup> *Planbureau voor de Leefomgeving* in Dutch and is the national institute for strategic policy analysis in the field of environment, nature and spatial planning. The distance to the nearest 100,000 jobs are given by a 500 by 500 meter cell. We combined the coordinates of these cells with the coordinates of the neighborhoods in the Amsterdam area and calculated the average distance for each neighborhood.

endogeneity of house prices are discussed in Section 2.2. Table 1 gives an overview of the data and its sources that are used in our analysis.

**Table 1.** Descriptive statistics household and neighborhood characteristics

Variables	Data source	Mean	SD	Min.	Max.
<b>Household characteristics</b>					
Gross primary household income	CBS (2008)	42,835	55,740	0	1,000,000
Household with children (-18)	CBS (2008)	0.240	0.427	0	1
Age of oldest household member	CBS (2008)	48.730	17.461	16	107
<i>Social Economic Category</i>					
Student	CBS (2008)	0.053	0.223	0	1
(Self-)Employed	CBS (2008)	0.559	0.496	0	1
Unemployed (Social assistance benefits)	CBS (2008)	0.176	0.381	0	1
Retired	CBS (2008)	0.212	0.409	0	1
<b>Neighborhood characteristics</b>					
Historic city center (km <sup>2</sup> )	RCE (2012)	0.027	0.134	0.000	1.029
Distance to the nearest 100,000 jobs (km)	PBL (2005)	8.287	3.355	0.637	18.407
Percentage rich households (%)	CBS (2008)	33.325	14.433	0.000	77.707
Price of standard house (in euros)	NVM (2009)	209,858	49,587	112,877	390,691

Note: We include 314 neighborhoods which covers most of the residential neighborhoods in the Amsterdam area. A few neighborhoods are left out because of the low number of household observations.

## 4 Estimation results

This section reports and discusses the results of the first and the second step of the sorting model for neighborhoods in the Amsterdam area. We provide an overview of the estimation results based on the basic specification, including the spatial spillovers of historic city centers, of the sorting model.

### 4.1 First step estimation results

In the first step of the residential sorting model, developed by Bayer, McMillan & Rueben (2004), we estimate the mean utilities (or: alternative specific constants) and the coefficients of Equation 2.6 via MNL with the location choice (neighborhood) of households as the dependent variable. The estimated coefficients are consistently estimated since we have no endogeneity issues on the household level, as we discussed in Section 2.2 (see footnote 8). Table 2 reports the

estimated coefficients,  $\beta_{kl}$ , of the cross effects between household and neighborhood characteristics of homeowners. The interpretation of the cross effect coefficients for income is straightforward. Since the household characteristics have a mean zero, the cross coefficients for income can be interpreted as the deviation from the mean utility (which is the utility of the average household). The interpretation of the cross effect coefficients for the social economic status are somewhat less straightforward because each household can only belong to one of the categories. Since we demeaned the dummy variables for the social economic status, this means we have to use all coefficients for each status to calculate the deviations from the mean indirect utility (see Equation 2.6 where in case of retired households,  $Z_{retired} = 1$  for retired households and  $Z_i = 0$  for the other categories and  $Z_i$  is the average of each category. Hence the estimated parameters of each category is needed to estimate the deviation of the mean indirect utility).

**Table 2.** First step results: Deviations from the alternative specific constant for homeowners

Neighborhood characteristics	Household characteristics		
	Income	Employed	Retired
Standardized house price (in euros)	<b>0.01259</b> (0.0006)***	<b>-1.3296</b> (0.0779)***	<b>2.2091</b> (0.0368)***
Historical city center (km <sup>2</sup> )	<b>0.00313</b> (0.0004)***	<b>-0.0967</b> (0.0526)	<b>0.0157</b> (0.0445)
Historical city center in surrounding neighborhoods	<b>-0.00031</b> (0.00005)***	<b>-0.0121</b> (0.0054)***	<b>0.0524</b> (0.0021)***
High income households (%)	<b>0.00027</b> (0.00001)***	<b>0.0233</b> (0.0059)***	<b>-0.1736</b> (0.0029)***
Distance to the nearest 100,000 jobs (km)	<b>0.00001</b> (0.00004)	<b>0.0310</b> (0.001)***	<b>0.0323</b> (0.0005)***

Note: Parameter estimates are used to calculate the deviations from the mean indirect utility with all variables normalized to have mean zero. Standard errors are in parentheses. Significance at 90%, 95% and 99% level are, respectively, indicated as \*, \*\*, and \*\*\*. The regression results based on other specifications can be obtained from the author.

The results give an indication how different types of households value amenities. For instance, results show that high income homeowners are less price sensitive than the average household. Homeowners that are retired seem to be less price sensitive than the employed. In other words, these types of homeowners are attracted to neighborhoods with a high standard house price. Most of the retired homeowners have benefit from the boom in house prices over the last decades. It is

most likely that they prefer to live in neighborhoods with a higher standard house price compared to the employed homeowners, where a large share did not benefit from the boom in house prices as much as the older generation.

We are, in particular, interested in the appreciation of the historic city center and the concentration of high income households. The results show that high income households prefer to live in neighborhoods that are within the historical city center. High income households are also attracted to a high concentration of high income households. These results imply that the historic city center attracts high income households. The sorting of these high income households increases the concentration of high income households, which further contributes to the attractiveness of the neighborhood for high income households. This suggests that there is a multiplier effect of the historic city center through attracting high income households. It is also likely that the increase in attractiveness of the neighborhood attracts many other (endogenous) amenities. This would suggest that the historic city center also has a multiplier effect through its impact on other consumer amenities, such as shops, restaurants, et cetera. We discuss this in the following section.

Employed and retired homeowners prefer to live in neighborhoods that are further from the large labor market and do not significantly differ in preferences for living in the historic city center. However, only the retired homeowners prefer to live close to the historic city center. The results suggest that retired homeowners, in contrary to employed homeowners, do not prefer to live in neighborhoods with a high share of high income households.

## 4.2 Second step estimation results

The second step of the residential sorting model consists of 2SLS estimation of Equation 2.7.<sup>9</sup> The dependent variable is the vector of mean indirect utilities – in other words that part of the utility that is equal for all households.<sup>10</sup> We deal with endogeneity through instrumental variables as discussed in Section 2.2. The instrument for house prices is computed as the equilibrium housing price that would prevail in the absence of unobserved heterogeneity. The instrument for the concentration of high income households is computed as the equilibrium concentration of high income households that would prevail in the absence of unobserved heterogeneity – simultaneously determined with the price instrument. The results of the estimations are reported in Table 3. Column 1 reports the simple OLS results. These coefficients suggest that the historic city

---

<sup>9</sup> First stage regression estimates of the 2SLS are available upon request. The null hypothesis of both under- and weak identification are rejected.

<sup>10</sup> The vector of mean indirect utilities,  $\delta_n$ , was estimated as alternative specific constants in the first step of the estimation procedure (Equation 2.6).

center is an important amenity for the location choice of the average household. However, the coefficients are likely biased as they do not account for the heterogeneity of prices and neighborhood socio-demographic characteristics. Column 2 reports the 2SLS results where only prices are instrumented. Some of the coefficients change substantially when we use the instrumental variables, notably the price coefficient, as is not uncommon in these models.<sup>11</sup>

**Table 3.** Second step estimation procedure: decomposition of the alternative specific constants

Variables	(1)	(2)	(3)
	OLS (se)	2SLS (se)	2SLS (se)
Standardized house price (in euros)	<b>-1.2582</b> (0.5621) **	<b>-26.6315</b> (7.976) ***	<b>-37.9354</b> (10.434) ***
Historical city center (km <sup>2</sup> )	<b>1.3146</b> (0.3482) ***	<b>5.7193</b> (1.9397) ***	<b>7.5236</b> (3.327) **
Historical city center in surrounding neighborhoods	<b>0.0521</b> (0.0435)	<b>1.2362</b> (0.3828) ***	<b>1.7907</b> (0.517) ***
High income households (%)	<b>-0.0079</b> (0.0087)	<b>0.1634</b> (0.0577) ***	<b>0.2618</b> (0.0812) ***
Distance to the nearest 100,000 jobs (km)	<b>-0.1323</b> (0.0285) ***	<b>-0.1383</b> (0.0922)	<b>-0.1692</b> (0.1393)
Constant	<b>15.5797</b> (6.661) **	<b>317.8204</b> (95.043) ***	<b>451.915</b> (124.15) ***
Price instrumented	No	Yes	Yes
High income households instrumented	No	No	Yes
F-statistic		11.427	6.598

Note: Standard errors are in parentheses. Significance at 90%, 95% and 99% level are, respectively, indicated as \*, \*\*, and \*\*\*. The first stage regression results can be obtained from the author.

Column 3 reports the 2SLS results where both prices and the concentration of high income households are instrumented. The coefficients change somewhat compared to Column 2. The coefficients have the expected sign and most of them are statistically significant. The results report a positive and significant effect of the historic city center and the spatial spillovers of the Amsterdam historic city center. This suggests that not only neighborhood inside the historic city center enjoy the benefits of the historic city center, but also those neighborhoods that are close to

<sup>11</sup> See, for example, Berry, Levinsohn & Pakes (1995).



the historic city center also benefit from this amenity. We thus find strong evidence that there are spatial spillovers of the historic city center. Column 3 also shows that the concentration of high income households has a positive and significant sign. This suggests that the attractiveness of a neighborhood also depends on the concentration of high income households. The results show that only the distance to the nearest 100,000 jobs is not statistically significant. This is probably because the research area is only the Amsterdam area. This area only covers one large labor market. There is probably an effect of the distance to a large labor market on the national level, but here we only take into account variation within the Amsterdam area.

## 5 Implications

In this section, we consider the implications of our estimation results of Section 4. We focus on the results in Column 3 of Table 3. The sorting model allows us to calculate the marginal willingness-to-pay (MWTP) of each type of household that we included in the analysis. These figures give a clear overview of the impact of different neighborhood characteristics on the location choice of heterogeneous households with respect to the price of a standard house. Furthermore, the sorting model also allows us to do a counterfactual analysis. The general equilibrium property, where housing demand has to match the housing supply, enables us to show how prices of a standard house change when we change one of the neighborhood or household characteristics. We report changes in the price of a standard house for several neighborhoods if there were no differences in the availability of the historic city center among all neighborhoods in the Amsterdam area.<sup>12</sup>

### 5.1 Marginal willingness-to-pay

The estimation results in Section 4 enable us to calculate the MWTP of heterogeneous households for neighborhood characteristics (see Appendix C for technical details). This allows us to compare the MWTP between neighborhoods. Column 1 of Table 4 reports the mean of the MWTP for all neighborhoods in the Amsterdam area. Columns 2 through 5 report the deviations from the mean for different types of homeowners.

The mean MWTP – in terms of higher prices for a standard house – for living inside the historic city center is large and significant (€41,619/km<sup>2</sup>). This implies that the price of the standard house increases on average with €41,619 for an extra square kilometer of historic city center. Since most neighborhoods are smaller than one square kilometer, this number decreases. The mean MWTP for

---

<sup>12</sup> Similar interpretations are the change in prices if there would be no historic city center in the Amsterdam area or if all households would not value the historic city center.

living not inside but close to the historic city center is also positive and significant (€9,906/km<sup>2</sup>). This number can be interpreted as an extra square kilometer of historic city center in surrounding neighborhoods – where the distance between adjacent neighborhoods is 1km (the average distance is somewhat lower than 1 km in the Amsterdam area) – that the average household is willing to pay in terms of the price for a standard house. The MWTP for living in neighborhoods with an extra percentage of high income households is positive and significant (€1,448/%), whereas the MWTP for living 1km closer to a large labor market is positive but it is not significant.

Column 2 reports the deviations from the mean MWTP for homeowners with different incomes. If the household earns €10,000 euro per year more than average, their MWTP for living inside the historic city center is around 5% higher than the average household. For living 1km from the historic city center, the MWTP is 1% higher than the average household. An extra percentage of high income households in the neighborhood, the MWTP is 10% higher than the average household. These numbers increase when the income of an household is even larger than €10,000 more than average.

**Table 4.** Marginal willingness to pay results from the 2SLS estimation

	(1) Mean	(2) Income (x10,000)	(3) Employed	(4) Retired
Historic city center (+km <sup>2</sup> )	40,274	2,175	-1,300 (ns)	1,838 (ns)
Historic city center in surrounding n'hoods (+km <sup>2</sup> )	9,842	91	-84	-193
High income households (+%)	1,414	137	55	161
Distance to nearest 100,000 jobs (-km)	644 (ns)	10 (ns)	-41	174

The values are in euros. (ns) means not significant at the 5% level. The significance levels of Column 2 to 4 are based on the first step estimation procedure of the residential sorting model.

The deviations from the mean MWTP for homeowners in different social economic categories are reported in Columns 3 and 4. The MWTP for employed and retired homeowners is different than the average household. Employed and retired homeowners are not willing to pay more for houses within or close to the historic city center. Employed and retired homeowners are willing to pay more for an extra percentage of high income households in the neighborhood. Employed homeowners are willing to pay more for houses closer to the labor market compared to the average household, whereas retired homeowners prefer to live further away from the labor market.

These results imply that retired homeowners tend to move away from the historic city center and away from neighborhoods close to a large labor market. These findings support our earlier

work (Van Duijn & Rouwendal, 2013) and are in line with the current literature on this topic (see, for example, Chen & Rosenthal, 2008).

## 5.2 Counterfactual analysis

The sorting model suggests that house prices react to changes in amenities. The general equilibrium property of the sorting model allows us to estimate the changes in house prices when the number of neighborhood amenities change. We have carried out a counterfactual simulation in which we compute the price of a standard house that would prevail if there were no differences in the availability of the historic city center in each neighborhood in the Amsterdam area. We set the area of historic city center at zero as if the historic city center would not exist. Evidently, the spatial spillovers of the historic city center will also disappear. This results in a new equilibrium and, therefore, new equilibrium prices. We then compute the price of a standard house for each neighborhood and scale them so the mean house price in the situation with and without the historic city center is identical. Because of the scaling, there will be neighborhoods where the price of a standard house will decrease (or increase) substantially.

**Table 5.** Counterfactual simulation: eliminating historic city center

Neighborhoods	Standardized house price (in euros)	Predicted house price (in euros)	Difference	Percentage
Amstel III en Bullewijk	119,581	191,581	+72,000	+60%
Bijlmer-Oost E, G en K	144,981	180,890	+35,909	+25%
Bijlmer-Centrum D, F en H	146,714	181,313	+34,599	+24%
Grachtengordel-Zuid	359,220	204,869	-154,351	-43%
Grachtengordel-West	359,694	204,790	-154,904	-43%
Museumkwartier	380,141	210,465	-169,676	-45%

Note: The predicted house prices, taken into account the general equilibrium property of the sorting framework and the scaling, are reported as a counterfactual simulation that sets all cultural heritage to zero.

Table 5 reports the prices of a standard house of the top 3 neighborhoods where the standard house price will increase and decrease (See Figure A.2 in Appendix A for the location of these neighborhoods). The simulation shows that some of the poorest neighborhoods in the Amsterdam area will benefit the most if the historic city center would not exist. The standard house prices in the most rich neighborhoods inside or close to the historic city center will decrease the most. Note

that the neighborhood *Museumkwartier* is just outside of the historic city center but, because of the large spatial spillover effects of the historic city center, house prices in this neighborhood will also decrease. The gap between the price of a standard house in these neighborhoods decreases, however, the prices in the city center will still be larger than in the other neighborhoods due to other amenities.

## 6 Multiplier effect

The results from Table 4 suggest that homeowners with a higher than average income are willing to pay more for houses within and around the historic city center. Consequently, this increases the share of high income households in the neighborhood. The results also suggest that homeowners with a higher than average income are willing to pay more for houses in neighborhoods with a large share of high income households. This suggests that urban heritage has a multiplier effect. We have a strong belief that the historic city center attracts many other (endogenous) amenities, which further magnifies the multiplier effect of urban heritage. In this section, we provide preliminary results that investigate the multiplier effects of the presence of a historic city center on other (endogenous) amenities.

The idea is that the number of amenities is explained by the presence of the historic city center, the share of high income households, population and other regressors on the neighborhood level. In Table 6 we report these results. The dependent variable in Column 1 is the number of grocery shops in each neighborhood, in Column 2 the number of fashion and luxury shops, and in Column 3 the number of leisure and culture shops. The results suggest that there are more shops in neighborhoods within the historic city center than in the other neighborhoods. There is no conclusive result that the number of shops is explained by the share of high income households in each neighborhood. The share of high income households in surrounding neighborhoods does also not give us a conclusive result that explains the number of shops. It seems the population of each neighborhood is an important driver that explains the number of shops. However, it is difficult to interpret these results because of the endogeneity issues involved in these simple regressions. The problem of causality is a difficult econometric issue to overcome in this type of research. However, with respect to the historic city center, we believe there is no causality problem. It is not possible (at least in the short term) to create a historic city center. This implies that the number of shops have no influence on the historic status of an area. Therefore, we believe that these results confirm that urban heritage has a multiplier effect on other (endogenous) amenities, in this case different

types of shops. In all the regressions, we find a positive effect of the presence of the historic city center on the number of shops.

**Table 6.** Simple regressions of different types of shops

	(1) Grocery shops (#) OLS (se)	(2) Fashion & Luxury shops (#) OLS (se)	(3) leisure & culture shops (#) OLS (se)
Historic city center (dummy)	15.0512 ** (6.9902)	54.3617 ** (24.2867)	125.7379 *** (36.5538)
Population (#)	0.0034 *** (0.0004)	0.0022 ** (0.0011)	0.0031 ** (0.0014)
High income households (%)	-0.0781 ** (0.0385)	0.01157 (0.1069)	-0.1896 (0.1237)
High income households in surrounding neighborhoods (%)	-0.1073 (0.0670)	-0.3684 * (0.2170)	-0.9158 *** (0.2982)
Standardized house price (in ln(euros))	6.8047 *** (2.1961)	18.3902 ** (7.7010)	30.5615 *** (7.0958)
Constant	-76.290 *** (26.339)	-211.388 ** (92.291)	-332.113 *** (84.581)
Observations	290	290	290
R-squared	0.6251	0.3091	0.5570

Note: Standard errors are in parentheses. Significance at 90%, 95% and 99% level are, respectively, indicated as \*, \*\*, and \*\*\*. These are preliminary results and more results (sensitivity analysis and control for neighborhood fixed effects) can be obtained from the author.

## 7 Conclusions

In this empirical paper, we investigate the location choice of households in the Amsterdam area. We compare households that differ in income, whether they are homeowners or renter, and to which social economic category they belong. We use a horizontal sorting model to measure the value that those heterogeneous households attach to a variety of amenities. The historic city center is such an amenity. We show that the impact of the historic city center on the location choice is large and significant. The total impact is the sum of its direct effect – living inside the historic city center improves the attractiveness of the neighborhood – and its indirect effect – neighborhoods inside the historic city center are attractive for other (endogenous) amenities. The results also suggest that spatial spillovers of the historic city center are present. Hence, neighborhoods that are just outside the historic city center are also attractive for households since they are still able to enjoy the amenities that are located within the historic city center.

Our analysis uses the sorting framework developed by Bayer, McMillan & Rueben (2004) in which the price of a standard house is explained by the housing supply and demand equilibrium.

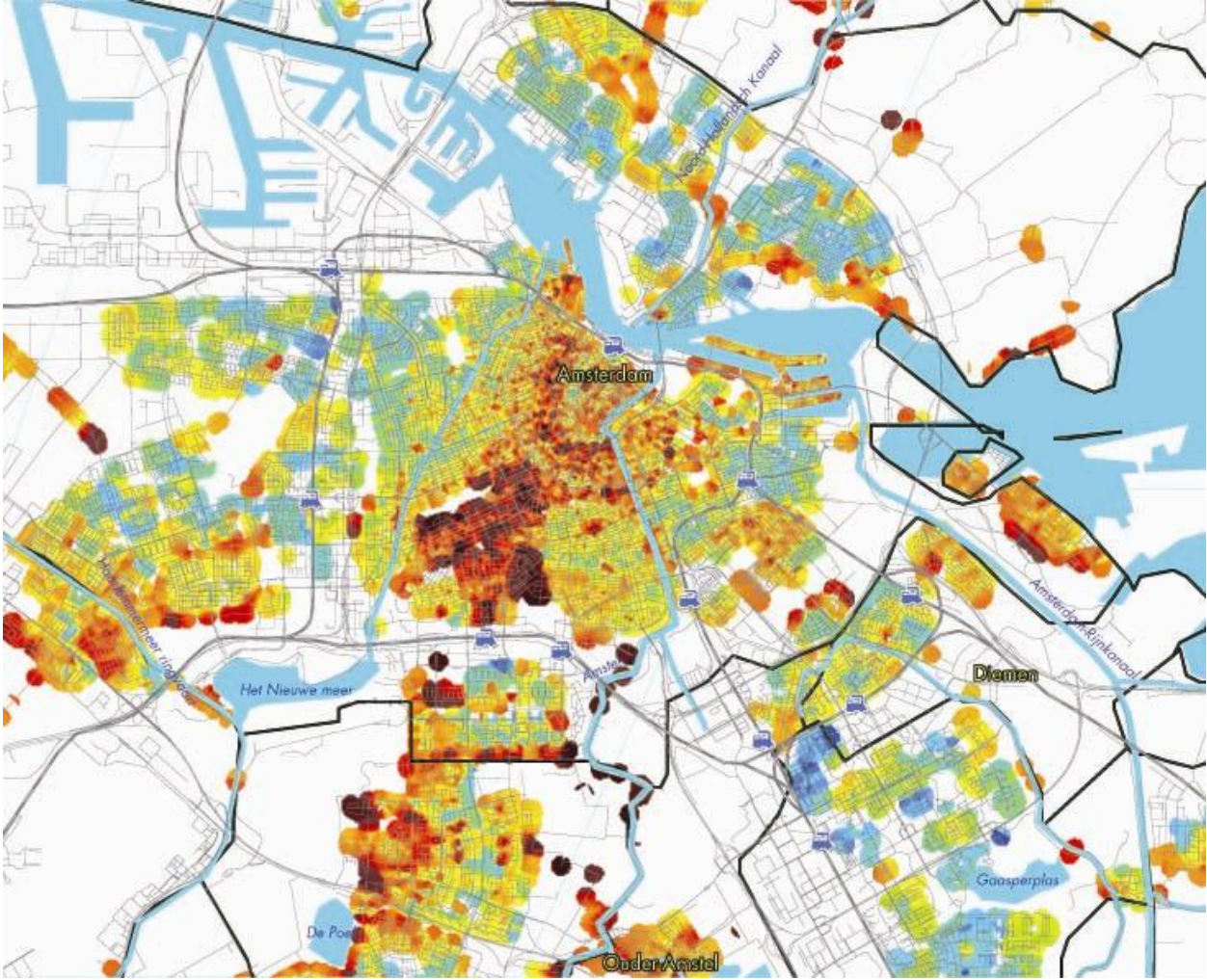
The idea is that different types of households reveal their preferences for neighborhood characteristics by choosing their neighborhood. The sorting framework allows us to calculate the marginal willingness-to-pay (MWTP) of different types of households for a variety of amenities. We show that for the Amsterdam area, the MWTP for living in a neighborhood inside the historic city center is highest for high income homeowners. We find that employed and retired homeowners are not different from the average household in their willingness to pay for living within the historic city center. The results show that employed homeowners prefer to live closer to the labor market, whereas retired homeowners prefer to live further from the labor market. Homeowners with a higher than average income are also attracted to neighborhoods with a large share of high income households. This suggests that the location choice of households in different stages of the life-cycle and with different incomes varies substantially.

In our counterfactual simulation, we show that if there would be no differences in the availability of the historic city center in each neighborhood in the Amsterdam area, the price of a standard house of neighborhoods inside or just outside the historic city center decreases by a large amount. These neighborhoods still have higher house prices than most other neighborhoods because of their favorable characteristics regarding their location close to other amenities.

In addition, we attempt to find some evidence that the historic city center has a multiplier effect on the attractiveness of neighborhoods. We believe that the historic city center attracts many other amenities. To prove the existence of this multiplier effect suffers one major problem: most of these amenities are endogenous. We include the concentration of high income households in our analysis. We control for the endogenous variables, price and concentration of high income households, using an instrumental variable strategy that takes advantage of the general equilibrium property of the sorting model. Our results suggest that high income homeowners are not only attracted to the historic city center but also to each other. This increases the concentration of high income households in the neighborhood, which further attracts more high income homeowners. We provide results of simple regressions that show some evidence of this multiplier effect. Neglecting the economic issues for the moment, we provide insights that show that the historic city center attracts many different types of shops while controlling for important factors like population and house prices. The share of high income households seem not to explain the number of shops in each neighborhood. These results are not conclusive and raise extra questions that are interesting for future research.



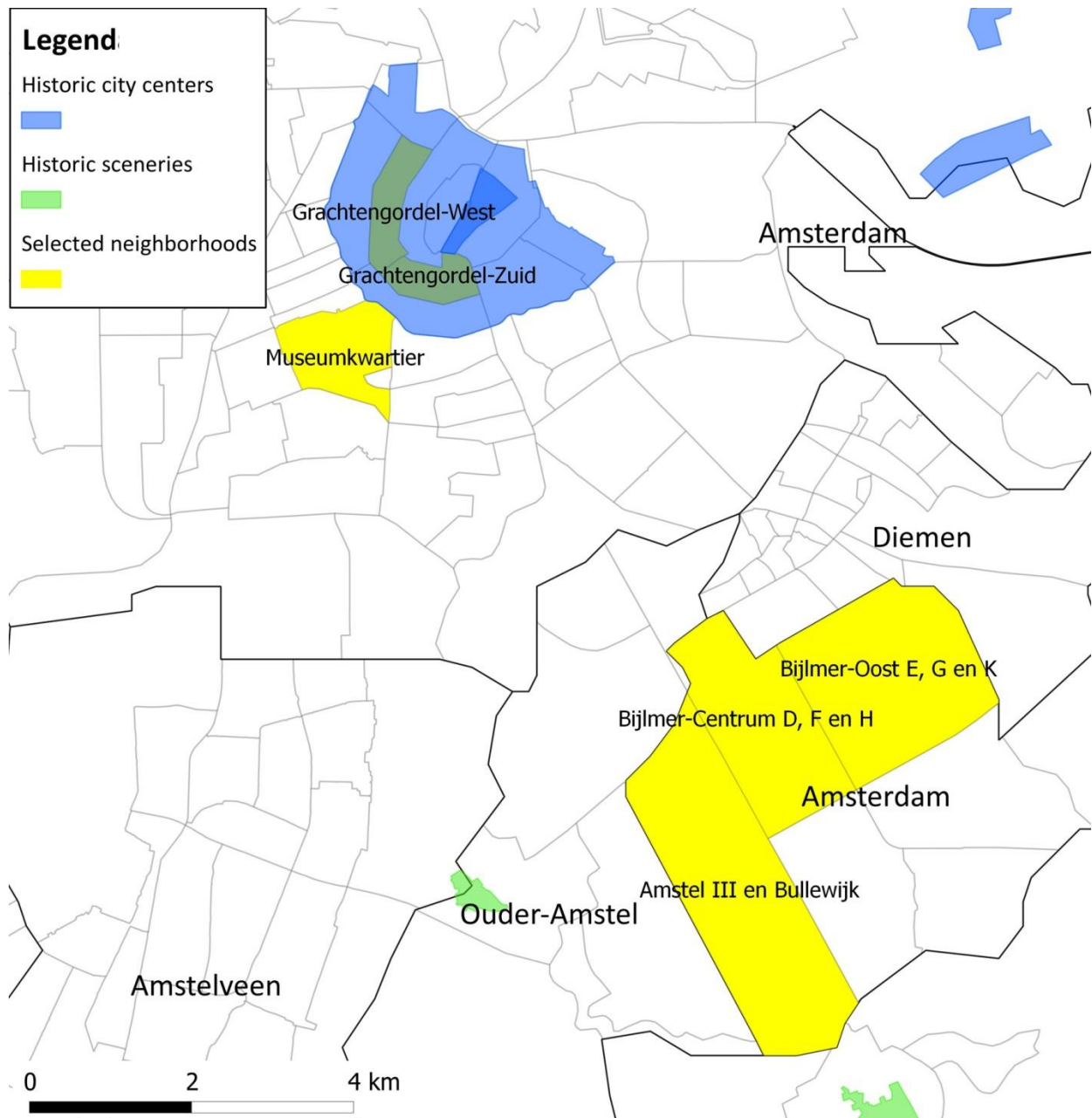
**APPENDIX A. MAPS OF THE AMSTERDAM AREA**



**Figure A.1.** Variation in house prices in Amsterdam.

Note: Dark red represents the highest house prices, green the lowest house prices. Prices are not controlled for structural characteristics.





**Figure A.2.** Neighborhoods in Amsterdam.

Note: The yellow areas are the top 3 neighborhoods that increased or decreased in the price of a standard house in the case when there is no differences in the availability of the historic city center (the blue areas).

## APPENDIX B. THE MULTIPLIER EFFECT

In this appendix we illustrate the effect of preferences with respect to demographic composition on the impact of a small change in cultural heritage for a simple two-good, two neighborhood example. The difference between the utilities of neighborhood 1 and 2 for group  $i$  ( $i = 1,2$ ) is:

$$\Delta u_i = \alpha_{0,i} + \alpha_{1,i} HP_1 - HP_2 + \alpha_{2,i} \sigma_1 - \sigma_2 + \alpha_{3,i} CH_1 - CH_2 .$$

And the choice probabilities are:

$$P_{1,i} = \frac{e^{\Delta u_i}}{1 + e^{\Delta u_i}}, \quad P_{2,i} = 1 - P_{1,i}.$$

In these equations  $HP$  denotes the housing price,  $\sigma$  the share of group 1 households and  $CH$  an indicator of cultural heritage, while the  $\alpha$ 's are coefficients.

We assume that group 1 is the rich group and that  $\alpha_{1,2} < \alpha_{1,1} < 0$ ,  $\alpha_{2,1} > \alpha_{2,2} > 0$ . That is, the price sensitivity of group 1 is less than that of group 2, whereas group 1 appreciates cultural heritage more than group 2. Assume for the moment that  $\alpha_{2,1} = \alpha_{2,2} = 0$ .

In equilibrium the housing market clears, so we must have:

$$N_1 P_{1,1} + N_2 P_{1,2} = S_1$$

where  $N_i$  is the number of households of group  $i$  and  $S_i$  is the number of houses in neighborhood  $i$ .

After a small change in cultural heritage in neighborhood 1, we must have:

$$N_1 dP_{1,1} + N_2 dP_{1,2} = 0$$

We can compute de changes in the choice probabilities as:

$$dP_{1,i} = P_{1,i} (1 - P_{1,i}) \alpha_{1,i} d\Delta HP + \alpha_{3,i} d\Delta CH .$$

Substitution of this result into the market equilibrium condition allows us to compute the change in the housing price difference as:

$$\frac{d\Delta HP}{d\Delta CH} = - \frac{N_1 P_{1,1} (1 - P_{1,1}) \alpha_{3,1} + N_2 P_{1,2} (1 - P_{1,2}) \alpha_{3,2}}{N_1 P_{1,1} (1 - P_{1,1}) \alpha_{1,1} + N_2 P_{1,2} (1 - P_{1,2}) \alpha_{1,2}}$$

Now consider the change in the share of rich households in neighborhood  $n$ ,  $\sigma_i$ . It equals:

$$\sigma_n = \frac{N_1 P_{n,1}}{S_n}$$

It can be verified that:

$$d\Delta\sigma = N_1 \left( \frac{1}{S_1} + \frac{1}{S_2} \right) dP_{1,1}$$

Since we have assumed that the rich are less sensitive to changes in the housing price and more sensitive to changes in cultural heritage, it must be the case that  $dP_{1,1} > 0$ . The change in the amount of cultural heritage therefore increases the share of rich households in neighborhood 1.

Let us now see what changes if we allow for preferences with respect to the demographic composition of neighborhoods. More specifically, assume that  $\alpha_{2,1} > 0$ ,  $\alpha_{2,1} > \alpha_{2,2}$ . That is, the rich have a preference to live among members of their own group and this preference is stronger than that of the poor to live among the rich. Note that we allow the poor to have positive as well as negative preferences to live among the rich.

The change in the probability  $P_{1,i}$  must now be computed as:

$$dP_{1,i} = P_{1,i} \left( 1 - P_{1,i} \right) \left( \alpha_{1,i} d\Delta HP + \alpha_{2,i} d\Delta\sigma + \alpha_{3,i} d\Delta CH \right)$$

This implies that for the change in  $\Delta\sigma$  we now have:

$$d\Delta\sigma = N_1 \left( \frac{1}{S_1} + \frac{1}{S_2} \right) P_{1,i} \left( 1 - P_{1,i} \right) \left( \alpha_{1,i} d\Delta HP + \alpha_{2,i} d\Delta\sigma + \alpha_{3,i} d\Delta CH \right)$$

Solving for  $d\Delta\sigma$  gives:

$$\begin{aligned} d\Delta\sigma &= \frac{N_1 \left( \frac{1}{S_1} + \frac{1}{S_2} \right) P_{1,1} \left( 1 - P_{1,1} \right) \left( \alpha_{1,1} d\Delta HP + \alpha_{3,1} d\Delta CH \right)}{1 - N_1 \left( \frac{1}{S_1} + \frac{1}{S_2} \right) P_{1,1} \left( 1 - P_{1,1} \right) \alpha_{2,1}} \\ &= M_1 N_1 \left( \frac{1}{S_1} + \frac{1}{S_2} \right) P_{1,1} \left( 1 - P_{1,1} \right) \left( \alpha_{1,1} d\Delta HP + \alpha_{3,1} d\Delta CH \right) \end{aligned}$$

where  $M_1 = 1 / \left( 1 - N_1 \left( \frac{1}{S_1} + \frac{1}{S_2} \right) P_{1,1} \left( 1 - P_{1,1} \right) \alpha_{2,1} \right)$ , a multiplier associated with the social interaction. Substituting this result in the expression we derived for the change in the choice probability  $P_{1,i}$ , we find:

$$dP_{1,i} = P_{1,i} (1 - P_{1,i}) \alpha_{1,i} d\Delta HP + \alpha_{3,i} d\Delta CH$$

$$+ \alpha_{2,i} M_1 N_1 \left( \frac{1}{S_1} + \frac{1}{S_2} \right) P_{1,1} (1 - P_{1,1}) \alpha_{1,1} d\Delta HP + \alpha_{3,1} d\Delta CH$$

The equation shows that the preferences for demographic composition result in an additional effect of the changes in the housing price and the cultural heritage on the choice probabilities.

## APPENDIX C. DERIVE THE MARGINAL WILLINGNESS-TO-PAY

Equation 2.6 and 2.7 can be written as a hedonic price regression allowing for heterogeneity in household preferences.

$$p_{k,n} = \frac{\beta_{0,k} + \beta_{k,l} Z_{i,l} - Z_l}{\beta_{0,p} + \beta_{p,l} Z_{i,l} - Z_l} X_{k,n} + \frac{1}{\beta_{0,p} + \beta_{p,l} Z_{i,l} - Z_l} \xi_n + \frac{1}{\beta_{0,p} + \beta_{p,l} Z_{i,l} - Z_l} \varepsilon_{i,n} - \frac{1}{\beta_{0,p} + \beta_{p,l} Z_{i,l} - Z_l} u_{i,n}$$

where  $p_{k,n} = \ln P_{k,n}$ . It is now simple to compute the MWTP of each  $i$  type of household for each neighborhood characteristic  $X_{k,n}$ :

$$\frac{\delta P_{k,n}}{\delta X_{k,n}} = \frac{\beta_{0,k} + \beta_{k,l} Z_{i,l} - Z_l}{\beta_{0,p} + \beta_{p,l} Z_{i,l} - Z_l} P_{k,n}$$

The household characteristics are constructed to have mean zero. This simplifies the MWTP of the average household:

$$\frac{\delta P_{k,n}}{\delta X_{k,n}} = \frac{\beta_{0,k}}{\beta_{0,p}} P_{k,n}$$

This is correct for calculation of the MWTP of the income variable. For the life-cycle status variables, the computation is somewhat more complicated. We divided the life-cycle status in four phases. Each head of the household is a student, (self-)employed, unemployed or retired. We used the (self-)employment status as a reference. For instance, when we want to calculate the MWTP of a retired household – note that the household characteristics have a mean zero – we have to incorporate the coefficients of the other categories as well. This is simply done by extending the MWTP computation of a retired household:

$$\frac{\delta P_{k,n}}{\delta X_{k,n}} = \frac{\beta_{0,k} + \beta_{k,retired} Z_{i,retired} - Z_l + \beta_{k,l} Z_{i,l} - Z_l}{\beta_{0,p} + \beta_{p,retired} Z_{i,retired} - Z_l + \beta_{p,l} Z_{i,l} - Z_l} P_{k,n}$$

where  $Z_{i,retired}$  equals one and  $Z_{i,l}$ , which represent the other categories, equals zero.