

Sorting based on urban heritage and income: Evidence from the Amsterdam metropolitan area

Revision: March 19, 2018

Mark van Duijn^{1,2,3}

Jan Rouwendal^{1,3,4}

¹ Department of Spatial Economics, VU University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

² Real Estate Centre, Population Research Center, University of Groningen, Landleven 1, 9747 AD Groningen, The Netherlands

³ Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, The Netherlands

⁴ Amsterdam School of Real Estate, Jollemanhof 5, 1019 GW Amsterdam, The Netherlands

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

Sorting based on urban heritage and income: Evidence from the Amsterdam metropolitan area

Mark van Duijn^{a,b,c} & Jan Rouwendal^{a,c,d}

^a Department of Spatial Economics, VU University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

^b Real Estate Centre, Population Research Center, University of Groningen, Landleven 1, 9747 AD Groningen, The Netherlands

^c Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, The Netherlands

^d Amsterdam School of Real Estate, Jollemanhof 5, 1019 GW Amsterdam, The Netherlands

Version: March 19, 2018

Abstract. Urban heritage is often concentrated in conservation areas with a protected status. Previous research argues that urban heritage attracts especially higher educated households who are likely to have higher incomes. The presence of these households may have a further impact on the attractiveness of the neighborhoods concerned, for instance through endogenous amenities like better shops or schools. If this is the case for high income households, conservation areas will have a further impact on the area's attractiveness through the demographic composition of the residential area. In this paper we investigate the interaction between the preference for urban heritage – as an exogenous amenity – and the preference for areas with a high concentration of high income households – as an endogenous amenity. We develop a logit-based sorting model in which different income groups interact and estimate it for the Amsterdam metropolitan area. Results show that all employed households highly value conservation areas and prefer to live in areas with a high concentration of high income households. We investigate the impact of urban heritage on house prices and welfare through counterfactual simulations. The disappearance of urban heritage would result in a substantially more suburbanized location pattern of the high income households in the Amsterdam metropolitan area, and to lower welfare for all income groups.

Key words: Location choice, urban heritage, sorting models, discrete choice, heterogeneous household preferences, welfare analysis

JEL classifications: R2, J1, Z1

Acknowledgements. The authors thank Kurt Schmidheiny, Stephan Heblich and the other participants of the Workshop on Cultural Heritage and Urban Revival, Nick Kuminoff and Lars Nesheim for useful comments on previous drafts. The authors also thank Statistics Netherlands, NVM, RCE and PBL for providing us with excellent data. The authors gratefully acknowledge Platform 31, VU University and the University of Groningen for their financial support.

1 Introduction

Urban amenities are closely related to the current urban revival. The decentralization of employment, the improved possibilities for communication and the secular decrease in transport costs have weakened the strength of traditional forces behind the concentration of economic activities. Since the amenities that make a city attractive for consumers, such as shops, theatres and restaurants remain localized 'consumer city' (Glaeser et al., 2001) becomes more important. Moreover, Brueckner et al. (1999) suggest that the presence of urban amenities is an important driver of the location of high and low income households in urban areas. In their theory, the demand for these amenities is highly income elastic. This implies that the rich will tend to locate in the city center when it is amenity-rich, while they will otherwise prefer the suburbs where land is much cheaper. In their theoretical model, amenities are taken as exogenous as is plausible for historic city centers. Other urban amenities, like shops, restaurants and theatres, are affected by the composition of the population in a neighborhood. Through this route, urban heritage may have a further impact on the attractiveness of neighborhoods. To the extent that this effect is positive, this secondary effect of the exogenous amenities reinforces their primary impact.

The economic literature on heritage has mainly focused on the impact of designation of monuments and conservation areas. See Navrud and Ready (2002), Noonan (2003) for surveys of the early literature. Examples are Coulson and Leichenko (2001), Coulson and Lahr (2005), and more recently Been et al. (2014), Ahlfeldt et al. (2017) and Koster and Rouwendal (2017). Recently there has been an increasing interest on cultural heritage in a wider urban economic context (Van Duijn and Rouwendal, 2013; Falck et al., 2015; Sheppard, 2015).

The relative location of high and low income households is an important issue in urban economics (see, for instance, Wheaton, 1977). The monocentric model, which is the workhorse of this literature, suggests that the ratio of the income elasticity of the demand for housing and that of the value of commuting time is the driving force of the spatial distribution of incomes in urban areas. If the value of time is roughly proportional to the wage (and hence income) and the income elasticity of housing demand is less than 1, as much of the literature suggests, the model predicts that the rich should live in the city center. Since many cities (in the US as well as elsewhere) do not confirm this prediction, other factors must be important. One possibility is the durability of housing which tends to make older housing, which is often overrepresented in the central city, less suitable

for high income households.¹ However, old houses are not necessarily inferior. For instance, in many European cities ancient buildings – the canal houses in Amsterdam are a clear example – are regarded as highly attractive urban heritage and are often inhabited by high income people.² This observation suggests that, under appropriate conditions, old housing may in fact contribute to the concentration of high income households in city centers.

Bayer et al. (2007) have documented for the San Francisco Bay Area that the attractiveness of neighborhoods for specific groups of households is influenced substantially by the presence of particular household groups.³ The pattern they observe confirms the sociological principle of ‘homophily’ – similarity breeds connection (see for instance, McPherson et al., 2001). There are several possible mechanisms behind this phenomenon. One is that households belonging to a particular group like to meet similar households and such interaction is facilitated by physical proximity. Another is that households with similar characteristics like the same type of amenities, such as schools, shops and restaurants with particular characteristics, and that these tend to emerge close to the concentrations of these households through the market or other allocation mechanisms as is suggested by the literature on Tiebout sorting.⁴ The preference of households to live in neighborhoods where similar households are located interacts with the exogenous amenities and may therefore reinforce their impact. In other words, an overrepresentation of high income households in conservation areas will have a further impact on the area’s attractiveness through preferences for demographic composition.

To investigate these issues, we develop and estimate a residential sorting model in which urban heritage – indicated by conservation areas – has a direct impact on the residential choice behavior of households, while we also incorporate preferences for demographic composition similar to Bayer et al. (2007) that may cause an additional impact of this amenity on neighborhood attractiveness for particular groups. Our model is of the ‘horizontal,’ logit-based type (see Kuminoff et al., 2013). This family of sorting models is closely related to the Berry, Levinsohn, and Pakes (1995) (BLP) model which is a workhorse model in industrial organization and was first applied to study residential sorting by Bayer and his co-authors.⁵

We make three methodological contributions to the literature on sorting models. First, we relax the assumption that all actors attach the same value to the unobserved characteristics of the

¹ See Bond and Coulson (1989) for an example of this ‘filtering’ literature, and Chen and Rosenthal (2008), and Brueckner and Rosenthal (2009) for recent contributions.

² Lazrak et al. (2014) find in a hedonic analysis on Dutch data that the prices of houses listed as monuments 15-20% higher than those of otherwise comparable houses.

³ The classical references for these sorting effects are Schelling (1971, 1978).

⁴ See Tiebout (1956) and Epple et al. (1984; 1998; 1999) for seminal contributions.

⁵ See Bayer, McMillan & Rueben (2004), Bayer & Timmins (2005; 2007), and Bayer & McMillan (2012).

choice alternatives. Second, we apply the methodology of Belloni, Chen, Chernozhukov, & Hansen (2012) to the choice of instruments for the price and the share of high income households in the sorting model. Third, we clarify the working of the instrument for the price used in Bayer, Ferreira, & McMillan (2007). These aspects of the paper are further discussed in section 2.

We find that households in all income groups attach a large value to urban heritage in their neighborhood, which we measure as the part of it included in a conservation area. The values we find are substantially larger than those reported in Van Duijn and Rouwendal (2013) who take the much larger municipalities as their spatial units of analysis.⁶ Our results also suggest that all households prefer to live in neighborhoods with a high share of high incomes. Although we do not find large differences in willingness to pay for urban heritage by income, a counterfactual simulation suggests that without urban heritage the location pattern of the high income households would be substantially different from what it is now, with significant concentrations of this group in some suburban locales, a pattern that is reminiscent of location patterns in U.S. metropolitan areas.

Our study sheds new light on the interaction between the effect of an exogenous amenity, cultural heritage, and that of the endogenous share of high income households on the attractiveness of neighborhoods. High income households are attracted to areas with cultural heritage, which makes these areas even more attractive to other households, which leads to a multiplier effect. Moreover, our simulations show that in the absence of cultural heritage in the city center, Amsterdam would look much more like a U.S. suburbanized metropolitan area, which confirms the idea that historical inner cities are an important background of the difference between European and American urban areas as was argued by Brueckner et al. (1999).

The remainder of this paper is structured as follows. Section 2 discusses the methodology concerning the residential sorting model, its issues, and the introduction of spatial elements in the model. Section 3 describes the data and discusses some descriptive statistics of household and neighborhood characteristics, which is followed by the estimation results in Section 4. In Section 5 we discuss the implications of the estimation results and Section 6 concludes.

⁶ Van Duijn and Rouwendal (2013) also develop a sorting model and estimate it on Dutch data, but there are a number of important differences with the present paper. They apply the model to all municipalities in the Netherlands, whereas we focus on the Amsterdam metropolitan area and use much smaller spatial units. They use survey data whereas the present paper uses administrative data. They do not consider the role of demographic composition, which is arguably less important at the municipal level, whereas this is a main focus of this paper. Moreover, they maintain the assumption that unobserved neighborhood characteristics have the same value for all households.

2 The location choice model

2.1 The first stage

Our methodology follows Berry, Levinsohn, and Pakes (1995) – from now on BLP – who addressed a number of important issues in discrete choice models of market demand. A main innovation of the BLP paper was that the possible presence of unobserved characteristics of the choice alternatives could be dealt with explicitly. This clarified an endogeneity issue associated with the price variable that had long plagued the estimation of logit models of market demand. BLP showed that, under appropriate assumptions, the endogeneity could be dealt with by including alternative-specific constants representing the average utility attached to the alternatives in the logit model. These alternative-specific constants absorb the impact of the unobserved characteristics as well as the endogenous price and they are analyzed further in a second estimation step using methods for linear equations.⁷

We assume that each consumer i ($i=1\dots I$) has preferences over n neighborhoods ($n=1\dots N$). These preferences refer to neighborhood characteristics that are not all observed by the researcher. Two observed characteristics that are of special importance are the neighborhood housing price, p_j , and the share of households belonging to a particular group in the neighborhood population, σ_j . This group is special because it evokes a ‘social interaction’ effect in that the choices of these actors affect the choice behavior of others in a direct way. We denote the other observed characteristics as a vector x and unobserved characteristics as a scalar ξ . Moreover we allow for idiosyncratic differences in preferences over alternatives ϵ .

The housing price and the share of households belonging to a particular group have a special status because they are determined by the model. The prices are pinned down by the market equilibrium condition that the share of actors choosing a choice alternative must equal the share of the housing stock present there. The choice behavior of the households belonging to the particular group affects the choice behavior of other groups in a direct way.

A restrictive assumption about the unobserved characteristics of the choice alternatives that, to the best of our knowledge, is imposed in all applications of BLP-type models is that they are evaluated identically by all actors. Since heterogeneity of preferences related to observed characteristics of the actors is generally found to be important in empirical work, this asymmetry is potentially restrictive. Below we generalize the BLP framework to a setup that allows a partial relaxation of this assumption by estimating separate models for subgroups of the population.

⁷ See Berry (1995) and BLP for further discussion. In these papers the price of the choice alternatives does not occur in the first (logit) stage of the estimation procedure.

That is, we assume that each consumer belongs to a group g ($g=1\dots G$) and that members of a given group have similar preferences. More specifically, we assume that the coefficients in the utility function referring to the price, the share of the special group and the unobserved neighborhood characteristics are identical for all members of a given group, while we allow the coefficients for the other neighborhood characteristics to be individual-specific. We specify the utility function for household i as:

$$\begin{aligned}
u_j^i &= \left[\alpha^{g(i)} \ln p_j + \beta^{g(i)} \sigma_j + \bar{\gamma}^{g(i)} x_j + \xi_j^{g(i)} \right] + \left[(\gamma^i - \bar{\gamma}^{g(i)}) x_j \right] + \epsilon_j^i \\
&= \delta_j^{g(i)} + \varphi_j^i + \epsilon_j^i,
\end{aligned} \tag{1}$$

where $g(i)$ denotes the group to which household i belongs, and $\delta_j^{g(i)}$ and φ_j^i in the second line are the first and second terms in square brackets, respectively, in the first line.

When $G=1$, our specification of the utility function is the conventional one in which δ_j equals the average utility attached to alternative j in the population and φ_j^i is the deviation of the deterministic part of the utility from the average. The obvious disadvantage of this specification is that it does not allow for heterogeneity in the evaluation of the endogenous variables p_j and σ_j and in the unobserved characteristics ξ_j . Although some researchers include such variables in the deviation from the average φ_j^i , a drawback is that the coefficients of this term have to be estimated in the first step where the endogeneity of these variables cannot be addressed. Opinions on the importance of this issue differ, but it seems desirable to avoid it if possible.

For $G>1$ some heterogeneity in the evaluation of endogenous and unobserved characteristics is allowed: heterogeneity across the groups is possible, but within each group homogeneity is still required. This may be regarded as a modest generalization of the conventional specification, but a potentially important one. For instance, in the setting of this paper it allows us to estimate different coefficients for the housing price and the share of the special group without evoking the concerns about endogeneity that are associated with introducing these variables in the first stage of the estimation procedure.

To estimate the model, we start with estimating G logit models with alternative-specific constants. For each of these models the sum of the estimated probabilities $\pi_j^i, i \in g$ that alternative j will be chosen by consumers in the group, divided by the total number of consumers in the group,

B^g , equals the observed share of households in that group choosing alternative j , which we denote as s_j^g ⁸:

$$\frac{\sum_{i:g(i)=h} \pi_j^i}{B^g} = s_j^g, \quad (2)$$

which implies that the contraction mapping technique can be used. It follows also that, for the total population we have that the sum of the estimated probabilities that alternative j will be chosen divided by the total population equal to observed share of the total population that chooses j :

$$\frac{\sum_i \pi_j^i}{B} = \frac{\sum_g B^g s_j^g}{B} = s_j. \quad (3)$$

This property will be used when computing the instrument for the price in the second stage.

Before moving to a discussion of that stage it is useful to observe that the procedure just proposed requires the presence of enough observations of actors belonging to each of the groups to be able to estimate the model. At the very least, there should be one observation of each group choosing each alternative, but in practice one would like to have more. This is not always an innocuous requirement. For instance, low and high income households may be strongly overrepresented in some parts of urban areas and strongly underrepresented in others. Extreme sorting would therefore invalidate the procedure just outlined. However, apart from this requirement, no additional assumptions are necessary to enable this method to work.

To complete the description of the model, we let H denote the set of actors belongs to the group causing social interaction. The shares σ_j are determined as:

$$\sigma_j = \frac{\sum_{i \in H} \pi_j^i}{B s_j}. \quad (4)$$

2.2 The second stage

The second stage consists of G linear regressions:

$$\delta_j^g = \alpha^g \ln p_j + \beta^g \sigma_j + \tilde{\gamma}^g x_j + \xi_j^g, \quad g = 1 \dots G, \quad (5)$$

where δ_j^g denotes the alternative specific constant for neighborhood j that has been estimated in the logit model referring to group g . The unobserved characteristics as evaluated by the various groups are the error terms in these equations. Since they may be correlated with the price and the share of high income households, these variables have to be instrumented.

⁸ This is a property of maximum likelihood estimation of the logit model as was shown by McFadden (1973).

Instruments have been difficult to find. The literature has therefore relied on the use of nonlinear functions of the exogenous characteristics of the choice alternatives, following the lead of BLP. These authors argued that the characteristics of other alternatives are excluded (by economic theory) from the utility of a given alternative and can therefore be used to construct instruments of the price. They proposed sums of the characteristics of sets of these alternatives. The intuition is that in a market with imperfect competition – like the automobile market - the price of a given alternative is set with an eye on the characteristics of its competitors. For the housing market a similar logic applies as the willingness to pay for a house at a given location is determined by the presence of good alternatives.

The econometric context of this issue is the fact that conditional moment restrictions imply the possibility to use arbitrary functions of exogenous variables as instruments. Chamberlain (1987) proposed the choice of instruments that minimize the asymptotic variance and BLP argue that their choice may be interpreted as an approximation to such optimal instruments. Approximations are useful, because exact derivation of the optimal instruments is often hard in practice.

The literature contains a number of different suggestions for functions of characteristics that can be considered as useful approximation to the optimal instruments. For instance, Bayer et al. (2007) noted, in the context of residential sorting, that the utility of living in a particular neighborhood can be affected by amenities in contiguous neighborhoods, which invalidates the exclusion restriction for using such amenities as instruments. They therefore exclude characteristics of choice alternatives that are close (within 3 km), and propose the use of characteristics that are nearby, but not too close as instruments (sometimes referred to as doughnut instruments).⁹ The literature in industrial organization following BLP contains several other suggestions for instruments. For instance, Dubé et al. (2011) use powers and cross products of the exogenous characteristics. The residential (Tiebout) sorting literature provides others. For instance, Epple and Sieg (1999) argue that the ranking of house prices per neighborhood is correlated with the ranking of an index of the local amenities. Indeed, the list of potential instruments one could construct from the exogenous characteristics of alternatives is infinitely large and in the absence of methods to compute the optimal ones, it remains arbitrary which ones to choose.

Belloni et al. (2012) have recently introduced a methodology for approaching optimal instruments through selection from a large number of potential candidates. Their basic assumption

⁹ This refers only to their first 'crude' instrument which they use as a first approximation. Their more refined instrument is discussed in the next subsection.

is that the model is approximately sparse, which means that the conditional expectation of the endogenous variable can be well approximated by a small number of candidate instruments. They show that variable selection by a particular variant of the lasso (= least absolute shrinkage and selection operator) is very useful for estimating the first stage of an IV regression. Belloni et al. (2012) furthermore show that the method allows for heteroscedasticity and present a procedure to deal with weak instruments. This method is potentially very useful as it offers the possibility to select the best of a number of potentially large instruments. Below, we use this method to select instruments from the large set of candidates that have been proposed in the literature.

2.3 Bayer's instruments

2.3.1 The housing price

In this subsection we pay attention to other instruments for the housing price and the share of a special group of households that have been proposed in the literature, but that use additional exogenous information. These instruments were developed in various papers with Bayer as a common co-author. Hence we refer to them as Bayer's instruments although other authors were involved in the various papers as well.

In the previous section we already mentioned Bayer et al. (2007)'s 'doughnut' version of the BLP instruments for the price associated with the choice alternatives. They use them for preliminary computations of the model and then switch to "*a more powerful instrument by calculating the predicted vector of market clearing prices for a version of the model that sets the vector of unobserved characteristics ξ to zero*" (Bayer, Ferreira and McMillan, 2007, p. 621). To understand how this instrument works, it is important to note that the heterogeneity of the population plays an essential role in its functioning. To see this, we consider the case of a single group ($G=1$), which is the setting in which Bayer et al. (2007) work. Moreover, for the moment we also ignore the 'social interaction' effect associated with the presence of a special group of households. Without heterogeneity the average utility of an alternative would be its actual utility for all actors and in the first stage we would only estimate the alternative-specific constants. Using (5), while suppressing the superscript g , we can write the housing price of alternative j as:

$$\ln p_j = (\delta_j - \gamma x_j - \xi_j)/\alpha. \quad (6)$$

The counterfactual price when the vector of unobserved characteristics ξ equals 0 is:

$$\ln p_j^{cf} (= \ln p_j + \xi_j/\alpha) = (\delta_j - \gamma x_j)/\alpha, \quad (7)$$

which would be useless as an instrument.

What happens if the population is heterogeneous? In Appendix B we derive two equations that are analogous to (6) and (7):¹⁰

$$\ln p_j = (\overline{EU^l} + \overline{\ln \pi_j^l} - \bar{\gamma} x_j - \xi_j) / \alpha. \quad (8)$$

$$\ln p_j^{cf} = (\overline{EU^{lcf}} + \overline{\ln \pi_j^{lcf}} - \bar{\gamma} x_j) / \alpha. \quad (9)$$

The first two terms in the braces on the right-hand side are the averages of the expected maximum utilities and the logged choice probabilities of alternative j in the population. Since demand for housing has to be equal to supply of housing, we must have $\overline{\pi_j^l} = s_j$, that is, the average of the probability that j will be chosen in the population must be equal to the share of the housing stock present in j . We can therefore rewrite $\overline{\ln \pi_j^l} = \frac{1}{B} \sum_i \ln \frac{\pi_j^l}{s_j} + \ln s_j$. The first term on the right-hand side is (apart from the sign) an entropy measure of inequality (see, for instance, Shorrocks, 1980, p. 622 or Theil, 1967, p. 126-7).

In the model, the heterogeneity in the choice probabilities is related to the heterogeneity of the characteristics of the actors. This distribution of the actor characteristics is exogenous information and thus can serve as the basis of an instrument. However, this information cannot be used in the way suggested by (8), because the choice probabilities depend on the unobserved characteristics. The suggestion of Bayer et al. (2007) is therefore to strip the unobserved characteristics from the model and compute the prices that will equilibrate demand and supply in that situation, given in (9).¹¹

Note that it is crucial to transform the information about the actor characteristics in such a way that it becomes alternative-specific. The differences in valuation of the characteristics of a choice alternatives determine the variation in choice behavior. Nonlinear interaction between the distribution of actor characteristics and the choice alternative characteristics is therefore crucial. The instrument uses the (counterfactual) choice probabilities for this purpose. The nonlinearity of discrete choice models is known to be a source of identification in itself, see Brock and Durlauf (2001), and this may contribute further to the usefulness of Bayer et al. (2007)'s instrument.

2.3.2 Related issues

The presence of a social interaction effect complicates the computation of the price instrument. If it would be included in the model when the counterfactual equilibrium prices are taken into account,

¹⁰ To verify the analogy, note that with homogeneous actors we have: $\pi_j = \exp(\delta_j) / \sum_k \exp(\delta_k)$, which implies that $\delta_j = \ln \pi_j + \ln \sum_k \exp(\delta_k)$ and note that $\ln \sum_k \exp(\delta_k)$, the 'logsum' gives the expected utility EU of the actors.

¹¹ See Bernasco et al. (2017) for the use of a similar instrument for criminal activity conditional upon neighborhood choice.

one can no longer maintain that the instrument is based only on exogenous information. This problem can be solved by leaving the social interaction effect out of the model when computing the price instrument. That is, β is set equal to 0 in this computation. Although this may lower the correlation between the instrument and the observed prices, the presence of all other characteristics will probably be sufficient to keep its value sufficiently high for the instrumental variable to work.

Another issue is that we also need an instrument for the shares σ_j of the actors belonging to the special group that causes the social interaction effect. We propose to construct it by, again, using the model with all ξ s set equal to zero and the prices set equal to the instrument values. That is, we compute the instrument values as the shares of the special group implied by the counterfactual situation in which there are no unobserved characteristics, prices set equal to their instrument values and the social interaction effect is also set equal to zero. That is, the instrument for the shares σ_j are the shares of the special group associated with the counterfactual situation to which our price instrument refers. This instrument thus uses the same information as the Bayer's price instrument, but in a different way. What we do is analogous to using a nonlinear function of one or more exogenous characteristics to find new candidate instruments.

The procedure for computing the two instruments can be refined somewhat through an iterative process. That is, one can substitute the instrument values for the σ_j s into the choice probabilities and set β at its appropriate value. This will distort the market equilibrium on which the price instrument has been based. Prices can then be re-adjusted to find a second counterfactual equilibrium in which the values of the instrument for the shares of the social interaction group have been taken into account. The implied choice probabilities can then be used to compute a second version of the instrument for these shares. And so on, until convergence. This is the procedure we used for computing Bayer's instruments as they have been used in the selection procedure discussed below.¹²

2.3 Other issues

In this subsection we deal with some remaining issues in the specification of our model. One is the specification of the heterogeneity in the evaluation of neighborhood characteristics by households belonging to the same group. A second concerns the possibility that some characteristics may have an impact that transcends the boundaries of the neighborhood in which they are located.

¹² Note that the computation of the instrument requires knowledge of the parameters in (7) and for that reason Bayer, McMillan and Rueben (2004) propose an iterative procedure that we adopt here. See Bayer, McMillan and Rueben (2004), p. 19-20.

The heterogeneity in preferences within groups g is embodied in the terms γ^i which are i -specific. The dimension of this vector is equal to the number of neighborhood characteristics K . We assume that its elements γ_k^i ($k = 1 \dots K$) are linear functions of the characteristics of the households. Denoting the value of characteristic l of household i as z_l^i , and postulate: $\gamma_k^i = \sum_{l=1}^L \omega_{k,l}^{g(i)} z_l^i$, where L is the number of household characteristics. Note that we allow this function to be specific for the group g to which the individual belongs. It follows then that we can write the second term in squared brackets on the right-hand side of (1) as:

$$(\gamma^i - \bar{\gamma}^{g(i)})x_j \left(= \sum_{k=1}^K (\gamma_k^i - \bar{\gamma}_k^{g(i)})x_{j,k} \right) = \sum_{k=1}^K \left(\sum_{l=1}^L \omega_{k,l}^{g(i)} (z_l^i - \bar{z}_l^{g(i)}) \right) x_{j,k}. \quad (10)$$

In the first stage of the estimation procedure we estimate the coefficients $\omega_{k,l}^g$ jointly with the alternative-specific constants δ^g for each group g . This provides a fairly flexible way for dealing with heterogeneity in household preferences.

Since our neighborhoods are relatively small, it is probable that their attractiveness is determined in part by amenities in other neighborhoods in the proximity. For instance, having the ancient Amsterdam city center within walking distance may still be experienced as an attractive property of a neighborhood, even though it does not belong itself to that center. We therefore allow for the possibility that some neighborhood characteristics are indicators of amenities in different, but close-by, neighborhoods. More specifically, if amenity k in neighborhood j affect the well-being of the inhabitants of surrounding neighborhoods, we define characteristic k' of a such a neighborhood j' as a distance-weighted average of amenity k in other neighborhoods:

$$x_{j',k'} = \sum_{j \in C_{j'}} e^{-\varphi_{k'} d_{j,j'}} x_{j,k}, \quad (11)$$

where $C_{j'}$ is the set of neighborhoods in the proximity of j' . This ‘potential’ formulation was also employed by Van Duijn and Rouwendal (2013) who used the much larger municipality as their spatial unit of analysis. Clearly, $x_{j',k'}$ in (11) can be interpreted as a spatial lag of $x_{j,k}$ with exponential weights.^{13, 14}

¹³ We only use the spatial lag for exogenous neighborhood characteristics.

¹⁴ The distance decay coefficient, φ , is set at 0.2. The function is therefore exponentially decreasing and weights are going towards zero when distance increases (weight < 0.1 if distance is 5 km). The cutoff point is set at 5km.

3 Data and descriptive statistics

3.1 Urban conservation areas

We estimate the model on microdata for the Amsterdam metropolitan area. The historic city center of Amsterdam and its canal belt are world-famous and have almost completely been listed as World Heritage by UNESCO in 2010. The center has many urban amenities like shops, restaurants, theatres, and has a cosmopolitan atmosphere which is regarded by many as very attractive. It is also a very popular residential area with high house prices. The ancient canal houses are still highly appreciated for residential purposes and only affordable by the rich.

In the Netherlands, the parts of urban areas that are regarded as being exceptionally valuable are listed as national conservation areas¹⁵ by the Netherlands Agency for Cultural Heritage,¹⁶ for their architectural and historic value. Apart from the historic city center several other parts of the Amsterdam metropolitan area have been given the status of a conservation area. The US equivalents of these conservation areas are the historic areas listed on the National Register of Historic Places under the authority of the National Park Service.¹⁷ The UK equivalents of listed buildings – determined on a national level – and conservation areas – designated by lower tier authorities – are listed by the Secretary of State for Culture, Media and Sport and based on the Planning Act 1990.

If the preferences of the households coincide (at least to a substantial extend) with the expert judgments that are behind the listing as conservations areas, one should expect that they are more attractive than non-listed areas. If the households that choose to live there are predominantly rich, they become overrepresented in these areas and if the composition of the neighborhood population is one of its relevant attributes for location choice, this will have a further impact on the housing market equilibrium in the area. Ignoring this effect when it is present may lead to biased estimates of the coefficients of a sorting model as the omitted variable is probably correlated with neighborhood characteristics that are included like urban heritage and house prices. The main purpose of our empirical work is to disentangle the effects of the urban heritage *per se* and that of the related sorting by income.

¹⁵ In Dutch: *Beschermde stadsgezichten*.

¹⁶ In Dutch: *Rijksdienst voor het Cultureel Erfgoed*. It is part of the Ministry of Education, Culture and Science.

¹⁷ 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 differs between countries.

3.2 The Amsterdam metropolitan area

The study area consists of the municipality of Amsterdam and a number of surrounding municipalities. The spatial unit we use is that of the neighborhood¹⁸ which is considerably smaller than a municipality. On average, around 1,800 households live in a single neighborhood. Comparing this to the average of the Netherlands, 630 households per neighborhood, it is clear that our study area focuses on the higher populated (urban) areas. In the study area there are 330 neighborhoods in 13 municipalities. We drop neighborhoods with none or only a few houses. These are mainly agricultural¹⁹ and industrial areas, such as Centrale Markt, Westelijk Havengebied and Bedrijventerrein Sloterdijk. Second, to keep the number of choice alternatives manageable and to focus on Amsterdam we aggregate most neighborhoods outside the Amsterdam municipality to the municipality level. This leaves us with 85 choice alternatives. Figure A.1 in Appendix A shows a map of the resulting areas. Below, we will still refer to these 85 choice alternatives as neighborhoods.

The boundaries of our (aggregated) neighborhoods do not always coincide with those of the conservation areas. We use the size of the area inside the boundaries of a neighborhood that belongs to a conservation area as an indicator of the amount of urban heritage in that neighborhood. For instance, the historic city center of Amsterdam is large, 679 hectare (6.79 km²),²⁰ and contains the canals, many gabled houses and numerous other monuments. There are ten neighborhoods that cover a part of it. The one with the largest part of the Amsterdam historic city center is *Nieuwmarkt en Lastage* (1.03 km²).

Figure 3 shows the share of the high income households, defined as the top 25% high income households, per neighborhood in the study area. The left panel refers to all households and it suggests that the rich are underrepresented in the city center. Although this may give the impression that Amsterdam is more similar to American cities than one would have expected, the background is entirely different. Housing policy in the Netherlands has for a long time emphasized the construction of social housing, especially in the large cities. Even in the center of Amsterdam, the share of social housing in the total stock is large, and since this sector is only accessible to low-income households, this has a substantial impact on the share of high income households in our study area.²¹ It is therefore also of interest to look at the share of high income households in the

¹⁸ In Dutch: *Buurt*. This is an administrative unit which does not only refer to residential areas.

¹⁹ Tight spatial planning in the Netherlands has resulted in the preservation of agricultural areas close to large cities. The 'Green Heart' of the Randstad is the primary example, but also to the north of Amsterdam the polder landscape is protected.

²⁰ The average size of conservation areas in the Netherlands is around 75 hectare (0.75 km²).

²¹ The social rental sector in Amsterdam covers approximately 50% of the housing stock – compared to the average of 35% in the Netherlands. Rents in the social sector are controlled according to a national system that links maximum possible rents to the number of quality points. The system ignores the typically large differences in house prices that are associated with urban location. As a consequence, houses with similar structural characteristics have similar rents everywhere in the Netherlands, which implies that social housing is especially cheap in the large urban areas like Amsterdam.

owner-occupied sector only, and this is what panel b) of Figure 3 does. Now the picture is different, and the share of high income households in some of the central area neighborhoods is higher than the surrounding neighborhoods.

As will be clear from this discussion, the owner-occupied and rental sectors of the Dutch housing market differ much. While the market mechanism determines allocation in the former part, rent control and the associated excess demand and queuing are dominant in the latter.²² We will take this difference into account when estimating the sorting model by treating rental and owner-occupied housing in a neighborhood as separate choice alternatives. This implies that the choice set has 170 alternatives: renting or owning in each of the 85 neighborhoods.

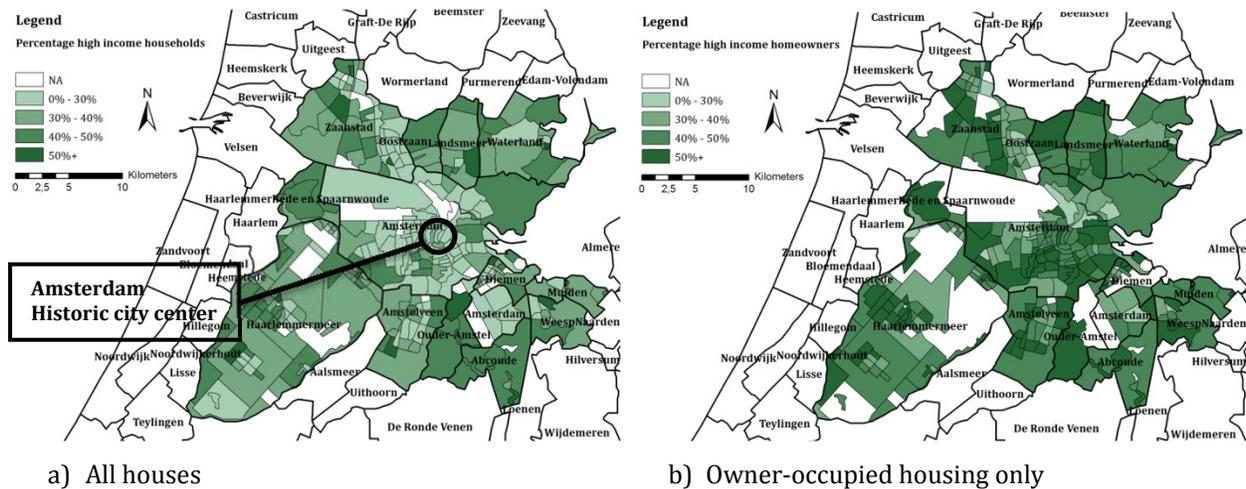


Figure 1. Distribution of high income households and high income homeowners in the Amsterdam area. Note: The figures are based on households within the highest income quartile (top 25%). Source: CBS (2008), own calculations.

3.3 Data

To estimate the sorting model we need information about households and neighborhoods. Administrative data on households was provided by Statistics Netherlands (CBS). The information refers to 2008 and contains approximately 600,000 households spread over the study area. We select households with at least one employed member and divide this population into three groups of equal size on the basis of income: high, middle and low income. For computational reasons, we take a random sample of these remaining 329,701 households based on the number of observations per choice alternative. For each choice alternative which contains 100 or more of each income group we take a random sample. This leaves us with a dataset in which we observe a minimum of

²² The allocation system for public housing is based on choice-based lettings. The private rental sector in the Netherlands is negligibly small.

30 households from each income group per choice alternative. In total this sample contains 86,663 households, which is 26% of the population of households in which at least one member is employed.²³

Table 1 shows an overview of the household and neighborhood characteristics. Household characteristics include income,²⁴ housing tenure, composition – whether the individuals within the household are a couple and whether the household has children under the age of 18 – age and neighborhood of residence. An important limitation of the CBS data is that they do not contain information about education levels. This is well known to be an important variable as it appears that especially young higher educated people appreciate urban heritage and the associated urban amenities. Although income and education are in general strongly correlated, this is less true for the younger households. Neighborhood characteristics include, apart from the conservation area, the price of a standard house, the share of high income households in the population of the neighborhood, distance to the nearest concentration of 100,000 jobs, distance to the nearest intercity train station, size of nature and water, and the average age of the housing stock. The price of a standard house is a hedonic price index based on transaction data from the Dutch Association of Real Estate Agents (NVM). We construct the index on the basis of a straightforward log-linear hedonic price regression that includes neighborhood dummies. This index gives the price of an owner-occupied house with a given bundle of characteristics. This enables us to compare prices within neighborhoods. The average of these prices is around €230,000.

Table 1. Descriptive statistics household and choice alternative characteristics

Variables	Data source	Mean	SD	Min.	Max.
Household characteristics (N=329,701)					
Gross primary household income (in €)	CBS (2008)	66,454	57,889	-	-
Couples	CBS (2008)	0.53	0.50		
Household with children (-18)	CBS (2008)	0.32	0.47		
Age of head of the household	CBS (2008)	42	11	-	-
Choice alternative characteristics (J=85)					
Conservation areas (km ²)	RCE (2012)	0.0991	0.2428	0	1.029
Percentage high income households (%)	CBS (2008)	32.56	12.35	-	-
Price of standard house (in €)	NVM (2009)	231,042	56,273	144,981	372,550
Distance to the nearest concentration of 100,000 jobs (km)	PBL (2005)	5.93	3.06	0.82	14.78
Distance to the nearest IC train station (km)	PBL (2005)	3.39	2.68	0.60	13.02
Nature (km ²)	CBS (2008)	0.94	1.36	0	7.63
Water (km ²)	CBS (2008)	0.68	1.02	0	4.65
Age of the housing stock	BAG (2010)	62	32	5	175

Note: Some values are hidden (-) because of confidentiality. See Figure A.1 of Appendix A for a complete picture of the 85 choice alternatives.

²³ This means that we have a choice-based sample. We use weighted maximum likelihood to deal with this characteristic (see Manski and Lerman, 1977). We used this sampling technique a few times to make sure that our results are robust.

²⁴ Statistics Netherlands provides information on gross primary household income.

4 Estimation results

This section reports and discusses the results of the two estimation steps of the sorting model for neighborhoods in the Amsterdam area. Because allocation of housing in the rental sector deviates from the market mechanism, we cannot regard the observed choices in this segment of the market as revealing the household's preferences in the same way as this occurs in a market setting. We have therefore estimated separate sets of coefficients for the owner-occupied and rental sectors and will only report estimation results of the former below.²⁵

4.1 First step estimation results

The first step of our residential sorting model involves the estimation of a MNL model for each of the three income groups. The utility function has been extensively discussed in Section 2 (see Equations 1 and 5). The utility that household i of group g attaches to alternative j is as follows:

$$u_j^i = \delta_j^{g(i)} + \sum_{k=1}^K \left(\sum_{l=1}^L \omega_{k,l}^{g(i)} (z_l^i - \bar{z}_l^{g(i)}) \right) x_{j,k}. \quad (12)$$

We estimate the alternative-specific constants $\delta_j^g, g = 1..G, j = 1 \dots J$ as well as the coefficients $\omega_{k,l}^g, g = 1 \dots G, k = 1 \dots K, l = 1 \dots L$. As discussed before, we do not include the possibly endogenous variables housing price and the share of high income households in the MNL model. Their impact will be absorbed by the δ_j^g 's. Also, separate sets of coefficients were estimated for alternatives referring to rental and owner-occupied housing.

Table 2 reports the estimated coefficients, $\omega_{k,l}^g$, for the owner-occupied market and the three income groups. Since the household characteristics are demeaned, these coefficients can be interpreted as referring to deviations from the mean utility of the group. The table shows a relatively large number of significant coefficients, indicating that there is non-negligible heterogeneity in tastes within the three income groups. For instance, results show that high income couples have a lower preference to live within conservation areas than the average high income household. Older high income households seem to have a higher preference to live within conservation areas than the average high income household. For high income households with children the effect is negative but not significantly different from zero. Note that the significant effect of couples on conservation areas disappears for the middle and low income households. This implies that especially high income singles have a positive preference to live within conservation areas that is significantly higher than that of the average high income household.

²⁵ The same procedure was used in Van Duijn and Rouwendal (2013).

Table 2. First step results for homeowners: Deviations from mean indirect utility

Neighborhood characteristics	Household characteristics					
	High income households					
	Couple		With children (-18)		Age	
Conservation areas (km2)	-0.404***	0.090	-0.145	0.091	0.057***	0.004
Conservation areas in surrounding neighborhoods	-0.117***	0.033	-0.060**	0.026	-0.010***	0.001
Distance to the nearest concentration of 100,000 jobs (km)	0.007	0.013	0.075***	0.010	0.005***	4.95E-04
Distance to the nearest intercity station (km)	0.092***	0.017	-0.033***	0.012	-0.002***	0.001
Nature (km2)	-0.023	0.022	0.066***	0.016	0.005***	0.001
Nature in surrounding neighborhoods	-0.062***	0.017	-0.032**	0.014	-0.001*	0.001
Water (km2)	0.125***	0.032	-0.015	0.022	0.004***	0.001
Water in surrounding neighborhoods	0.035*	0.021	-0.001	0.016	0.003***	0.001
Age of the neighborhood	-0.002**	0.001	0.001	0.001	1.43E-04***	4.16E-05
Log-likelihood	-133,928					
	Middle income households					
	Couple		With children (-18)		Age	
Conservation areas (km2)	-0.049	0.118	-0.167	0.142	0.055***	0.005
Conservation areas in surrounding neighborhoods	-0.140***	0.035	-0.146***	0.037	8.15E-05	0.001
Distance to the nearest concentration of 100,000 jobs (km)	-0.007	0.013	0.023*	0.014	0.003***	0.001
Distance to the nearest intercity station (km)	0.089***	0.016	0.030*	0.016	0.001	0.001
Nature (km2)	0.000	0.023	0.024	0.024	0.006***	0.001
Nature in surrounding neighborhoods	-0.064***	0.019	-0.040*	0.021	0.003***	0.001
Water (km2)	0.096***	0.031	0.000	0.031	0.004***	0.001
Water in surrounding neighborhoods	0.027	0.020	0.068***	0.021	-0.003***	0.001
Age of the neighborhood	0.002	0.001	-0.003**	0.001	-1.78E-05	5.04E-05
Log-likelihood	-137,213					
	Low income households					
	Couple		With children (-18)		Age	
Conservation areas (km2)	-0.069	0.145	-0.271	0.169	0.050***	0.005
Conservation areas in surrounding neighborhoods	-0.134***	0.046	-0.019	0.050	3.37E-04	0.002
Distance to the nearest concentration of 100,000 jobs (km)	0.007	0.018	0.025	0.019	0.001*	0.001
Distance to the nearest intercity station (km)	-0.006	0.022	0.030	0.023	0.003***	0.001
Nature (km2)	-0.019	0.029	0.025	0.032	0.004***	0.001
Nature in surrounding neighborhoods	-0.077***	0.026	0.006	0.029	0.003***	0.001
Water (km2)	0.069*	0.041	0.059	0.044	0.006***	0.002
Water in surrounding neighborhoods	-0.006	0.028	0.004	0.031	-0.003**	0.001
Age of the neighborhood	0.002	0.002	-0.005***	0.002	1.68E-04***	6.02E-05
Log-likelihood	-135,997					

Note: Parameter estimates are used to calculate the deviations from the mean indirect utility with all variables normalized to have mean zero. Significance at 90%, 95% and 99% level are, respectively, indicated as *, **, and ***. The first step results for renters can be found in Table D.1 of Appendix D. The regression results based on other specifications can be obtained from the author.

4.2 Second step estimation results

The second step of the residential sorting model consists of three 2SLS estimations based on Equation 5. The dependent variable is the vector of mean indirect utilities for each income group δ_j^g . There are two endogenous variables: the logged price and the share of high income households. We used the procedure suggested in Belloni et al. (2012) to find the best among the following sets of instruments:

- The squares and cross products of the exogenous characteristics
- The cubes and third-order cross products of the exogenous characteristics
- BLP-like instruments, including the ‘doughnut’ versions
- Rankings of the characteristics
- Bayer’s instruments as discussed in 2.3

These instruments and the results of LASSO are further discussed in Appendix C.²⁶ We initially limited the set of instruments first to these groups to see how the various suggestions for candidate instruments that have been made in the literature work out for our data. We also provide results for the complete set of candidate instruments. Since we have two endogenous variables, the price and the share of high income households, we carried out two instrument selection procedures. In general this resulted in the choice of different instruments. In the second stage we used the union of the sets of the selected instruments for the two variables as instruments.

Table 3 reports the results of the second stage for the three groups of households. We test for weak instruments by using a Wald F statistic based on the Kleibergen-Paap rk statistic (Kleibergen and Paap, 2006). The results tell a clear story: despite selection of the best instruments from the various groups as suggested by Belloni et al. (2012), with the single exception of Bayer’s instruments the selected instruments are weak. Only in the case of BLP-like instruments and Bayer’s instruments the estimated coefficients for the price and the share of high income households are significant. If we select from the complete set of candidate instruments, only the two Bayer instruments are chosen (see column 6). If we use the set of all instruments except the two Bayer instruments, we still have weak instruments. Therefore, we only discuss the results of Column (6).

²⁶ We used the STATA code provided on Christian Hansen’s research page for applying LASSO.

Table 3. Second step results (1 of 3): Mean indirect utility for each income group

	High income households						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Squares and cross products	Cubes and third order cross product s	BLP-like instrument s	Rank instrument s	Bayer's instrument s	All candidate instrument s	All except Bayer's instrument s
Ln(Price of a standard house)	0.622 (2.736)	-63.29 (150.6)	-5.903** (2.815)	4.557 (4.752)	-10.02*** (2.018)	-10.02*** (2.018)	-3.684 (3.712)
Share high income households (%)	0.0631 (0.0607)	0.303 (2.798)	0.145*** (0.0462)	0.121 (0.0893)	0.116*** (0.0186)	0.116*** (0.0186)	0.116*** (0.0449)
Conservation areas	-0.603 (0.538)	11.96 (37.22)	0.0576 (0.508)	-2.159 (1.766)	1.352** (0.533)	1.352** (0.533)	-0.157 (0.542)
Conservation areas in surrounding neighborhoods	-0.0701 (0.208)	5.060 (12.02)	0.415* (0.249)	-0.435 (0.494)	0.775*** (0.212)	0.775*** (0.212)	0.250 (0.306)
Distance to the nearest concentration of 100,000 jobs (km)	-0.0624 (0.0543)	-0.224 (0.603)	-0.087 (0.0608)	-0.0626 (0.113)	-0.0911 (0.0556)	-0.0911 (0.0556)	-0.0785 (0.0548)
Distance to the nearest IC train station (km)	0.00774 (0.136)	-0.373 (5.182)	-0.137 (0.126)	-0.102 (0.239)	-0.079 (0.0791)	-0.079 (0.0791)	-0.0862 (0.117)
Nature (km2)	0.0391 (0.112)	0.724 (4.694)	0.0104 (0.176)	-0.127 (0.294)	0.131 (0.179)	0.131 (0.179)	0.0218 (0.140)
Nature in surrounding neighborhoods	-0.131 (0.120)	1.782 (8.049)	-0.0976 (0.119)	-0.452 (0.392)	0.152* (0.0922)	0.152* (0.0922)	-0.106 (0.105)
Water (km2)	-0.0704 (0.181)	0.742 (8.609)	-0.172 (0.18)	-0.353 (0.414)	0.0241 (0.167)	0.0241 (0.167)	-0.135 (0.159)
Water in surrounding neighborhoods	0.0735 (0.105)	-2.112 (7.585)	-0.0121 (0.0956)	0.381 (0.339)	-0.260*** (0.094)	-0.260*** (0.094)	0.0147 (0.0972)
Age of the neighborhood	-0.00860* (0.0052)	0.130 (0.385)	-9.62E-05 (0.00685)	-0.0242 (0.0188)	0.0132* (0.00759)	0.0132* (0.00759)	-0.0029 (0.00708)
Constant	-8.649 (32.25)	749.8 (1783)	68.29** (33.06)	-55.96 (56.61)	117.5*** (23.9)	117.5*** (23.9)	42.13 (43.83)
Observations	85	85	85	85	85	85	85
Weak F-test	0.556	0.022	2.48	0.297	23.929	23.929	2.095

Note: Dependent variable is δ_j^g which represents the alternative specific constants estimated in the first step of the sorting model. Standard errors are in parentheses. Significance at 90%, 95% and 99% level are, respectively, indicated as *, **, and ***.

Table 3. Second step results (2 of 3): Mean indirect utility for each income group

	Middle income households						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Squares and cross products	Cubes and third order cross product s	BLP-like instru- ments	Rank instru- ments	Bayer's instru- ments	All candidate instru- ments	All except Bayer's instru- ments
Ln(Price of a standard house)	-0.459 (2.620)	-54.53 (125.3)	-6.698** (2.640)	4.044 (4.188)	-10.76*** (1.906)	-10.76*** (1.906)	-2.443 (1.838)
Share high income households (%)	0.0259 (0.0585)	0.250 (2.309)	0.109** (0.0433)	0.065 (0.0797)	0.0856*** (0.0175)	0.0856*** (0.0175)	0.0427* (0.0245)
Conservation areas	-0.578 (0.520)	9.816 (31.16)	-0.0044 (0.498)	-2.071 (1.513)	1.219** (0.524)	1.219** (0.524)	-0.289 (0.36)
Conservation areas in surrounding neighborhoods	-0.0512 (0.194)	4.274 (10.03)	0.409* (0.228)	-0.451 (0.429)	0.761*** (0.202)	0.761*** (0.202)	0.102 (0.174)
Distance to the nearest concentration of 100,000 jobs (km)	-0.0732 (0.050)	-0.213 (0.498)	-0.0975* (0.058)	-0.0697 (0.0996)	-0.102* (0.0545)	-0.102* (0.0545)	-0.0795* (0.0453)
Distance to the nearest IC train station (km)	0.028 (0.128)	-0.335 (4.276)	-0.121 (0.116)	-0.0487 (0.211)	-0.0724 (0.0757)	-0.0724 (0.0757)	-0.00114 (0.069)
Nature (km2)	0.0486 (0.103)	0.59 (3.911)	0.0119 (0.165)	-0.0959 (0.255)	0.123 (0.172)	0.123 (0.172)	0.0537 (0.0924)
Nature in surrounding neighborhoods	-0.123 (0.116)	1.433 (6.716)	-0.106 (0.111)	-0.416 (0.339)	0.126 (0.0908)	0.126 (0.0908)	-0.090 (0.0801)
Water (km2)	-0.0649 (0.167)	0.552 (7.136)	-0.180 (0.161)	-0.302 (0.365)	-0.00194 (0.157)	-0.00194 (0.157)	-0.0698 (0.132)
Water in surrounding neighborhoods	0.0398 (0.103)	-1.757 (6.338)	-0.0293 (0.0936)	0.328 (0.294)	-0.262*** (0.0894)	-0.262*** (0.0894)	-0.00569 (0.0693)
Age of the neighborhood	-0.00677 (0.00476)	0.108 (0.323)	8.40E-04 (0.00673)	-0.022 (0.016)	0.0135* (0.00762)	0.0135* (0.00762)	-0.00339 (0.00398)
Constant	6.661 (30.90)	648.1 (1484)	80.18*** (31.03)	-47.25 (49.93)	128.8*** (22.55)	128.8*** (22.55)	30.13 (21.80)
Observations	85	85	85	85	85	85	85
Weak F-test	0.556	0.022	2.48	0.297	23.929	23.929	2.457

Note: Dependent variable is δ_j^g which represents the alternative specific constants estimated in the first step of the sorting model. Standard errors are in parentheses. Significance at 90%, 95% and 99% level are, respectively, indicated as *, **, and ***.

Table 3. Second step results (3 of 3): Mean indirect utility for each income group

	Low income households						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Squares and cross products	Cubes and third order cross products	BLP-like instruments	Rank instruments	Bayer's instruments	All candidate instruments	All except Bayer's instruments
Ln(Price of a standard house)	0.174 (2.588)	-56.55 (128.4)	-6.135** (2.608)	3.489 (3.93)	-10.19*** (1.924)	-10.19*** (1.924)	-2.114 (1.863)
Share high income households (%)	0.0089 (0.0559)	0.279 (2.335)	0.0921** (0.0407)	0.0599 (0.0774)	0.0752*** (0.0179)	0.0752*** (0.0179)	0.0302 (0.0241)
Conservation areas	-0.389 (0.499)	10.14 (32.08)	0.206 (0.504)	-1.729 (1.459)	1.350** (0.545)	1.350** (0.545)	-0.0778 (0.374)
Conservation areas in surrounding neighborhoods	-0.0997 (0.200)	4.414 (10.28)	0.366 (0.223)	-0.409 (0.400)	0.713*** (0.199)	0.713*** (0.199)	0.0752 (0.179)
Distance to the nearest concentration of 100,000 jobs (km)	-0.0487 (0.0494)	-0.201 (0.507)	-0.0731 (0.0573)	-0.0492 (0.0968)	-0.0788 (0.0556)	-0.0788 (0.0556)	-0.0563 (0.046)
Distance to the nearest IC train station (km)	-0.00341 (0.127)	-0.449 (4.325)	-0.151 (0.115)	-0.101 (0.205)	-0.116 (0.0803)	-0.116 (0.0803)	-0.0407 (0.0684)
Nature (km ²)	0.0667 (0.102)	0.574 (4.004)	0.032 (0.163)	-0.0778 (0.244)	0.13 (0.172)	0.13 (0.172)	0.0692 (0.0951)
Nature in surrounding neighborhoods	-0.133 (0.119)	1.40 (6.88)	-0.113 (0.11)	-0.412 (0.323)	0.099 (0.092)	0.099 (0.092)	-0.101 (0.0809)
Water (km ²)	-0.032 (0.164)	0.502 (7.25)	-0.143 (0.159)	-0.278 (0.358)	0.0112 (0.159)	0.0112 (0.159)	-0.044 (0.133)
Water in surrounding neighborhoods	0.0464 (0.102)	-1.755 (6.504)	-0.0267 (0.0911)	0.312 (0.279)	-0.242*** (0.0884)	-0.242*** (0.0884)	-0.00122 (0.0688)
Age of the neighborhood	-0.00402 (0.00446)	0.113 (0.332)	0.00381 (0.00663)	-0.0174 (0.0154)	0.0157** (0.00765)	0.0157** (0.00765)	-3.22E-04 (0.00413)
Constant	-0.601 (30.47)	672 (1521)	73.75** (30.64)	-40.48 (46.82)	122.1*** (22.76)	122.1*** (22.76)	26.44 (22.03)
Observations	85	85	85	85	85	85	85
Weak F-test	0.556	0.022	2.48	0.297	23.929	23.929	2.457

Note: Dependent variable is δ_i^j which represents the alternative specific constants estimated in the first step of the sorting model. Standard errors are in parentheses. Significance at 90%, 95% and 99% level are, respectively, indicated as *, **, and ***.

The results in Column (6) of Table 3 show that, for all income groups, households appear to appreciate living close to high income households. The coefficient is significantly higher for high income households than for low income households. This result confirms that of Bayer et al. (2007) who found self-segregation based on income and ethnicity while all households prefer to live close to high income households. This suggests it is more likely that low income households are pushed out of gentrifying neighborhoods, rather than leaving them because they dislike their changing demographic composition.

The coefficients of conservation areas are positive and significant for all income groups. Somewhat unexpected is that the coefficients between income groups are not significantly different from each other. A priori, we expected that the average high income household would have a

higher preference to live close to conservation areas than the average medium or low income household. This result suggests that income might not affect the preference of living close to conservation areas, but, again, that high income households are likely to push out the other income groups because they can spend more on housing. However, results from the first step discussed in Section 4.1 suggest that other household characteristics, such as marital status and age, are also important in determining the preference of conservation areas.

We observe similar results for the spatial lag of conservation areas. The positive and significant coefficients show that conservation areas have a positive external effect on the attractiveness of locations just outside the neighborhoods in which they are located. This result is in line with Ahlfeldt et al. (2017) and Been et al. (2017). It shows that spatially lagged exogenous variables are important to include in location choice models. Households do not only enjoy the characteristics from their own neighborhood but also those from surrounding neighborhoods.

Only few of the other neighborhood characteristics that we include in our location choice model are also important for the location choice of households. The coefficients for the distance to the nearest concentration of 100,000 jobs and for the distance to the nearest intercity train station show a negative but insignificant effect. This is not too surprising as the Amsterdam area can be regarded as one labor market with good accessibility to public transport. For nature we find the expected positive signs of the estimated coefficients, but they are not significant. For water, we find inconclusive results for the main effect but a negative and highly significant effect for its spatial lag. Both amenities are relatively abundantly present in the suburban areas constructed in the postwar period that are less appreciated and we therefore suspect that these variables pick up the presence of the associated negative amenities.²⁷ The age of the housing stock should be expected to absorb this effect to some extent. The coefficients have a positive sign, indicating that older neighborhoods are in general more attractive. This is in contrast to most US city centers where low income households seem to occupy old homes (Rosenthal, 2008).

5 Implications

In this section, we consider the implications of our estimation results reported in Section 4. 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

²⁷ In their hedonic price analysis Rouwendal and Weijsschede-van der Straaten (2008) also found that the green areas in the suburban quarters of Amsterdam to have no impact on house prices.

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.²⁸ We report changes in the price of a standard house for several areas within the Amsterdam metropolitan area if there were no differences in the availability of conservation areas among all neighborhoods in the Amsterdam area.²⁹

5.1 Marginal willingness-to-pay

The estimation results reported in Section 4 enable us to calculate the MWTP of heterogeneous households for neighborhood characteristics (see Appendix E for technical details). This allows us to compare the MWTP between the income groups and neighborhoods. Column 1 of Table 4 reports the average MWTP of the three income groups. For the high income group the figures represent the MWTP of the average high income household. Columns 2 through 4 report the deviations from the average of each income group for couples, households with children and age.

The mean MWTP – in terms of higher prices for a standard house – for an additional percentage point of high income households in their neighborhood is largest for high income households (€2,666 for a percentage point increase). As we discussed above, we did not allow for differences in the MWTP within the three income groups to prevent endogeneity issues. The differences between the groups that we find are small. Low income households are still willing to pay around €1,700 to live in a neighborhood with a 1% higher share of high income households.

The MWTP of an average high income household for living inside conservation areas is large and significant (€31,183 for an additional km² in conservation area). This implies that the average high income household, that has to pay €31,181 extra for a house in an area with an extra square kilometer of conservation area in their neighborhood, still reaches the same utility as when living in an otherwise equal neighborhood without this amenity.³⁰ The figures for the other income groups are somewhat lower but of the same order of magnitude. The figures are much higher than those reported in Van Duijn and Rouwendal (2013) who used much larger spatial units. Their results refer to living in the municipality Amsterdam instead of another municipality, whereas the results reported here allow us to differentiate between various neighborhoods within the municipality

²⁸ Note that our model does only explain relative house prices. We therefore assume that the average house price remains unchanged in counterfactual simulations.

²⁹ Similar interpretations are the change in prices if there would be no conservation areas in the Amsterdam area or if all households would not value conservation areas.

³⁰ Note that the ceteris paribus condition involved refers also to the random part of the utility function.

Amsterdam. The results thus show that there is substantial heterogeneity within the municipal boundaries, as seems plausible.

Table 4. Marginal willingness to pay results

High income households	Average household	Deviations from the mean		
		Couple	Household has children (-18)	Age (+10 years)
High income households (+%)	2666 ***			
Historic city center (+km ²)	31183 ***	-1986 ***	-1977	13068 ***
Historic city center in surrounding neighborhoods	9637 ***	-310 ***	-440 **	-1231 ***
Distance to the nearest concentration of 100,000 jobs (+km)	-2101	34	1024 ***	1041 ***
Distance to the nearest intercity station (+km)	-1822	453 ***	-453 ***	-515 ***
Nature (+km ²)	3029	-114	901 ***	1101 ***
Nature in surrounding neighborhoods	1890 *	-165 ***	-236 **	-152 *
Water (+km ²)	555	614 ***	-205	941 ***
Water in surrounding neighborhoods	-3233 ***	93 *	-5	324 ***
Age of the neighborhood (+year)	304 **	-12 **	13	33 ***
Middle income households				
High income households (%)	1837 ***			
Historic city center (km ²)	26168 ***	-534	-2511	11843 ***
Historic city center in surrounding neighborhoods	8810 ***	-821 ***	-1182 ***	9
Distance to the nearest concentration of 100,000 jobs (+km)	-2195 *	-73	344 *	543 ***
Distance to the nearest intercity station (km)	-1555	964 ***	453 *	194
Nature (km ²)	2633	-3	367	1358 ***
Nature in surrounding neighborhoods	1455	-377 ***	-324 *	397 ***
Water (km ²)	-42	1047 ***	-3	836 ***
Water in surrounding neighborhoods	-3032 ***	159	552 ***	-324 ***
Age of the neighborhood (+year)	289 **	21	-44 **	-4
Low income households				
High income households (%)	1705 ***			
Historic city center (km ²)	30618 ***	-1064	-4596	11244 ***
Historic city center in surrounding neighborhoods	8718 ***	-1115 ***	-173	41
Distance to the nearest concentration of 100,000 jobs (+km)	-1788	110	432	272 *
Distance to the nearest intercity station (km)	-2630	-87	506	704 ***
Nature (km ²)	2958	-296	422	875 ***
Nature in surrounding neighborhoods	1211	-641 ***	58	364 ***
Water (km ²)	254	1066 *	998	1413 ***
Water in surrounding neighborhoods	-2959 ***	-53	38	-336 **
Age of the neighborhood (+year)	356 **	33	-83 ***	38 ***

Note: The values are in euros. Significance at 90%, 95% and 99% level are, respectively, indicated as *, **, and ***. The significance levels of Column 2 to 4 (deviations from the mean) are based on the first step estimation procedure of the residential sorting model.

Living not inside but close to conservation areas is also contributes significantly to well-being. The mean MWTP is highest for high income households (€9,637). This number can be interpreted as the value attached to an extra square kilometer of conservation area in surrounding neighborhoods – where the distance between adjacent neighborhoods is 1km (the average distance between the cores of neighborhoods in the Amsterdam area is somewhat lower than 1 km) – that the average household is willing to pay in terms of the price for a standard house.

The interpretation of the mean MWTP figures of the other neighborhood characteristics is similar. Deviations from the mean of each group are shown in Columns 2 through 4 of Table 4. Their

interpretation can be clarified as follows: If a household belongs to the high income group, and is a couple with children under 18 while the age of the head of the household is equal to the average in this group, their MWTP for an additional square kilometer of conservation area is €3,963 ($= -1986 + -1977$) lower than the average high income household. This MWTP is around 10% lower than that of the average high income household.

These results show that there is large heterogeneity between different types of households. On average, high income couples seem to prefer areas outside the historic center and further from the intercity station where the housing stock is younger compared to the average high income household. Middle income couples seem only to prefer to live further from the intercity station and in areas with water compared to the average middle income household. Low income couples seem to not differ that much from the average low income households. High income households with young children prefer to live in neighborhoods outside conservation areas with more green, further away from the labor market but close to an intercity station compared to the average high income household. Middle and low income households with young children prefer younger neighborhoods compared to the average household of that group. The age component seems to play a large role as well. Note that we did not include retired households in the sample so the positive deviations from the mean for living in conservation areas is not surprising. Depending to the income group, households where the head of the household is 10 years older than the average, their MWTP for living in conservation areas increases with around €11,000 to €13,000 in terms of house prices (which is around a 35% increase from the mean MWTP). The same positive, but slightly smaller, numbers are found for the presence of nature and water. This suggests that older households who are still working have a higher preference to reside in areas within conservation areas with preferably a lot of green and water compared to younger households.

Some of the effects we find are not monotone in income. Low income households are sometimes more similar to high incomes households than medium income households. This may be related to our lack of information about education that was noted earlier and the attractiveness of Amsterdam for highly educated young people, who often do not (yet) have a high income. More general it may have to do with the attractiveness that a cosmopolitan city like Amsterdam has on specific groups of people who like the special atmosphere of the city although living there does not enable them to reach a particularly high income.

5.2 Impact on house prices

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 conservation areas in each neighborhood in the Amsterdam area. We set conservation areas at zero in all neighborhoods.³¹ Evidently, the spatial spillovers of conservation areas will also disappear but all other neighborhood characteristics remain unchanged.

Table 5. Counterfactual simulation: Set the conservation areas at zero in all neighborhoods

Municipalities – Districts	Standardized house price (in €)	Predicted house price (in €)	Difference in prices	High income households (in %)	Predicted high income households (in %)
Abcoude	237,154	349,149	111,995	59.0	80.0
Amstelveen	227,539	283,357	55,818	45.3	48.4
Amsterdam:					
Stadsdeel Centrum	310,047	223,399	-86,648	32.1	24.2
Stadsdeel Noord	183,164	190,361	7,197	24.7	22.5
Stadsdeel West	249,824	214,506	-35,318	23.5	18.6
Stadsdeel Nieuw-West	183,086	215,425	32,339	27.5	28.7
Stadsdeel Zuid	296,512	271,588	-24,924	37.2	35.1
Stadsdeel Oost	232,712	221,535	-11,177	29.5	28.8
Stadsdeel Zuidoost	156,902	190,791	33,889	20.2	19.3
Diemen	192,874	226,353	33,478	39.9	40.2
Haarlemmerliede	195,849	253,658	57,809	47.5	55.1
Haarlemmermeer	188,879	262,474	73,595	46.6	60.1
Landsmeer	206,782	310,377	103,594	52.1	75.2
Muiden	224,802	267,977	43,175	50.1	58.3
Oostzaan	189,357	269,137	79,781	48.6	66.1
Ouder-Amstel	222,688	286,674	63,986	51.4	58.1
Waterland	194,735	265,102	70,367	49.1	66.6
Weesp	204,458	217,276	12,817	36.1	29.3
Zaanstad	156,280	199,865	43,585	37.2	42.6

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 urban heritage to zero.

We use the following procedure. First, we compute the new price equilibrium while keeping the shares of high income households constant in the demand equations.³² At this price equilibrium

³¹ Results will not change if we set the variable at any other value, e.g. the average over all neighborhoods. The important thing is that there are no differences in this amenity between the neighborhoods.

³² The price equilibrium computed in this way is unique (see, for instance, Rouwendal, 1990).

the share of high income households will differ from what they were originally. The second step is to adjust these shares and re-compute the price equilibrium. This procedure is continued until the shares of high income households do no longer change. In each step the house prices are scaled so as to keep the average housing price equal to its value in the situation where urban heritage was present.³³

Table 5 reports the results of the simulation where we set the conservation areas at zero in all neighborhoods. For a better overview of the effects we aggregate the choice alternatives to the district³⁴ level for Amsterdam and to the municipal level for the other neighborhoods. Column 1 reports the average price of a standard house – determined by a simple hedonic model with neighborhood fixed effects. Column 2 reports the adjusted price of a standard house after we set the conservation areas at zero in all neighborhoods. This is the counterfactual price that would prevail if there was no urban heritage after the adjustment of the share of high income households. Column 3 reports the difference between Column 1 and 2. As expected, the central areas of Amsterdam become much cheaper, while the price in suburban districts of Amsterdam and the surrounding municipalities increases. The standard house prices in the rich Amsterdam districts which are within or close to the historic city center decrease the most. Note that the district *Amsterdam Stadsdeel Zuid* is 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 Amsterdam and elsewhere decreases. However, the prices in the city center will still be higher than in most other areas in Amsterdam due to the attractive central location close to the Central station and the accessibility to the job market.

Columns 4 and 5 of Table 5 report the actual percentage of high income households and the counterfactual percentage of high income households after we set the area of conservation areas at zero in all neighborhoods. For many districts and municipalities the share of the high income households hardly changes. However, we find strong increases in the share of high income households in suburban municipalities where the initial share of high income households was already high, such as Abcoude and Landsmeer, and almost no changes in Diemen and Amstelveen, which are closer to the city of Amsterdam. This suggests that when urban heritage in the city center disappears, high income households are likely to move to ‘posh’ suburban municipalities. Our simulations thus confirm the impression that urban heritage contributes significantly to keep the rich households – especially those rich workers who are single, older and without young children as can be concluded from the MWTP figures – in the center of the metropolitan area. These results are

³³ The procedure used by Bayer and McMillan (2012) is identical.

³⁴ In Dutch: *Wijk of stadsdeel*. This is an administrative unit which is between the neighborhood and municipal level.

in line with the amenity-based theory of Brueckner et al. (1999) that high income households will only locate in city centers if city centers have a strong amenity advantage over suburban municipalities.

6 Conclusion

In this empirical paper, we use a logit-based sorting model to investigate the impact of urban heritage on demographic composition and house prices in a metropolitan region. Our analysis uses the sorting framework developed by BLP and Bayer et al. (2007) in which the price of a standard house is explained by the housing supply and demand equilibrium. To mitigate the concerns about endogeneity, we extended the sorting model by splitting the population into three groups, based on income, and estimated different logit models for each of these groups. Within each group we allow for heterogeneous preferences of different households. In particular, we allow for differences in the sensitivity for the housing price and the share of high income households between these groups, but not within them.

Another novel aspect of our analysis is the use of Belloni et al. (2012)'s selection procedure for optimal instruments. We find that all candidate instruments – that are nonlinear functions of the neighborhood characteristics and do not use any additional exogenous information – perform poorly. In contrast, the instruments suggested by Bayer and his co-authors in several papers perform very well. We show that these instruments use additional information that combines information about household characteristics – crossed with the neighborhood characteristics – and the nonlinearity of the model.

We find that all households attach a large value to the proximity of urban heritage and that differences in proximity to heritage within the municipality are important. We also find that households attach value to the share of high income households living in a neighborhood. All households prefer to live in neighborhoods with a higher share of high income households, which is in line with Bayer et al. (2007). We find that all income groups attach roughly the same value to urban heritage, but the value varies with other household characteristics. This does not confirm our a priori expectations based on Brueckner et al. (1999) – that high income households attach a larger value to the proximity of urban heritage than lower income households – but it unravels new insights in the value of urban heritage.

We use the estimated model to calculate marginal willingness to pay for neighborhood amenities and to compute counterfactual equilibrium prices for the situation in which urban heritage would be absent. We find a large willingness to pay for urban heritage that contributes to

substantial differences in attractiveness and, as a consequence, in house prices between areas in the Amsterdam metropolitan region. Although we found relatively weak evidence for Brueckner et al.'s (1999) suggestion that especially high income households are attracted to the urban amenities, our simulations show that the disappearance of urban heritage would lead to a substantially more suburbanized location pattern of the high income households in our study region.

References

- Ahlfeldt, G.M., K. Möller, S. Waights, and N. Wendland (2017): "Game of zones: The political economy of conservation areas", *Economic Journal*, 127(602), F421-F445.
- Bayer, P., F.V. Ferreira, and R. McMillan (2005): "Tiebout Sorting, Social Multipliers and the Demand for School Quality", *NBER Working Paper No. 10871*.
- Bayer, P., F.V. Ferreira, and R. McMillan (2007): "A Unified Framework for Measuring Preferences for Schools and Neighborhoods", *Journal of Political Economy*, 115, 588-638.
- Bayer, P., R. McMillan, and K. Rueben (2004): "An Equilibrium Sorting Model of Sorting in an Urban Housing Market", *NBER Working Paper No. 10865*.
- Bayer, P., and C. Timmins (2005): "On the Equilibrium Properties of Locational Sorting Models", *Journal of Urban Economics*, 57, 462-477.
- Bayer, P., and C. Timmins (2007): "Estimating Equilibrium Models of Sorting across Locations", *The Economic Journal*, 117, 353-374.
- Bayer, P. and R. McMillan (2012) Tiebout sorting and neighborhood stratification. *Journal of Public Economics*, **96**, 1129-1143.
- Been, V., I. Gould Ellen, M. Gedal, E. Glaeser, B.J. McCabe (2016): "Preserving history or restricting development? The heterogeneous effects of historic districts on local housing markets in New York City", *Journal of Urban Economics*, 92, 16-30.
- Bernasco, W., Th. De Graaff, J. Rouwendal and W. Steenbeek (2017): "Social Interactions and Crime Revisited: An Investigation Using Individual Offender Data", *Review of Economics and Statistics*, 99, 622-636.
- Berry, S. (1994): "Estimating Discrete-Choice Models of Product Differentiation", *The RAND Journal of Economics*, 25, 242-262.
- Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium", *Econometrica*, 63, 841-890.
- Bond, E.W., and N.E. Coulson (1989): "Externalities, Filtering, and Neighborhood Change", *Journal of Urban Economics*, 26(2), 231-249.
- Brueckner, J.K., and S.S. Rosenthal (2009): "Gentrification and Neighborhood Housing Cycles: Will America's Future Downtowns Be Rich?", *The Review of Economics and Statistics*, 91(4), 725-743.
- Brueckner, J.K., J.F. Thisse, and Y. Zenou (1999): "Why is Central Paris Rich and Downtown Detroit Poor? An Amenity-based Theory", *European Economic Review*, 43, 91-107.
- Chen, Y., and S.S. Rosenthal (2008): "Local Amenities and Life-cycle Migration: Do People Move for Jobs or Fun?" *Journal of Urban Economics*, 64, 519-537.

- Epple, D., R. Filimon, and T. Romer (1984): "Equilibrium Among Local Jurisdictions: Towards an Integrated Approach of Voting and Residential Choice", *Journal of Public Economics*, 24, 281-304.
- Epple, D. and G. J. Platt (1998): "Equilibrium and Local Redistribution in an Urban Economy When Households Differ in Both Preferences and Income", *Journal of Urban Economics*, 43, 23-51.
- Epple, D., and H. Sieg (1999): "Estimating Equilibrium Models of Local Jurisdictions", *Journal of Political Economy*, 107, 645-681.
- Falck, O., M. Fritsch, S. Heblich, and A. Otto (2015): "Music in the Air: Estimating the Social Return to Cultural Amenities", *Working paper*.
- Kleibergen, F., and R. Paap (2006): "Generalized reduced rank tests using the singular-value decomposition", *Journal of Econometrics*, 127, 97-126.
- Koster, H.R.A, and J. Rouwendal (2017): "Historic Amenities and Housing Externalities: Evidence from the Netherlands", *Economic Journal*, 127(605), F396-F420.
- Lazrak, F., P. Nijkamp, P. Rietveld, and J. Rouwendal (2014): "The Market Value of Cultural Heritage in Urban Areas: An Application of Spatial Hedonic Pricing", *Journal of Geographical Systems*, 16(1), 89-114.
- Manski, C.F. and S.R. Lerman (1977) "The estimation of choice probabilities from choice-based samples" *Econometrica*, 45, 1977-1988.
- McPherson, M., L. Smith-Lovin, and J.M. Cook (2001): "Birds of feather: Homophily in social networks", *Annual Review of Sociology*, 27, 415-444.
- McFadden, D. (1973): "Conditional Logit Analysis of Qualitative Choice Behaviour", in ed. P. Zarembka, *Frontiers of Econometrics*. Academic Press, New York.
- Rosenthal, S.S. (2008): "Old homes, externalities, and poor neighborhoods. A model of urban decline and renewal", *Journal of Urban Economics*, 63(3), 816-840.
- Rouwendal, J. (1990): "Existence and Uniqueness of Stochastic Price Equilibria in Heterogeneous Markets", *Regional Science and Urban Economics*, 21, 218-242.
- Rouwendal, J., and W. Weijschede-van der Straaten (2008): "The Costs and Benefits of Providing Open Space in Cities", *CPB Discussion Paper 98*.
- Schelling, T.C. (1971): "Dynamic Models of Segregation", *Journal of Mathematical Sociology*, 1, 143-186.
- Schelling, T.C. (1978): *Micromotives and Macrobehavior*. New York: Norton.
- Sheppard, S. (2015): "Culture Concentrations: The spatial structure of arts nonprofits", *Working paper*.
- Shorrocks, A.F. (1980): "The Class of Additively Decomposable Welfare Measures", *Econometrica*, 48, 613-625.
- Theil, H. (1967): *Economics and Information Theory*. Amsterdam: North Holland.
- Tiebout, C.M. (1956): "A Pure Theory of Local Expenditures", *Journal of Political Economy*, 64, 416-424.
- Van Duijn, M., and J. Rouwendal (2013): "Cultural Heritage and the Location Choice of Dutch Households in a Residential Sorting Model", *Journal of Economic Geography*, 13(3), 473-500.
- Weijschede-Van der Straaten, W., J. Rouwendal, and J. Möhlmann (2017): "Why are commuting distances of highly educated couples so short? An analysis of the heterogeneity of the location preferences of households", *Working paper*.

Wheaton, W. (1977): "Income and Urban Residence: An Analysis of Consumer Demand for Location", *American Economic Review*, 620-632.

APPENDIX A. OVERVIEW OF THE CHOICE ALTERNATIVES

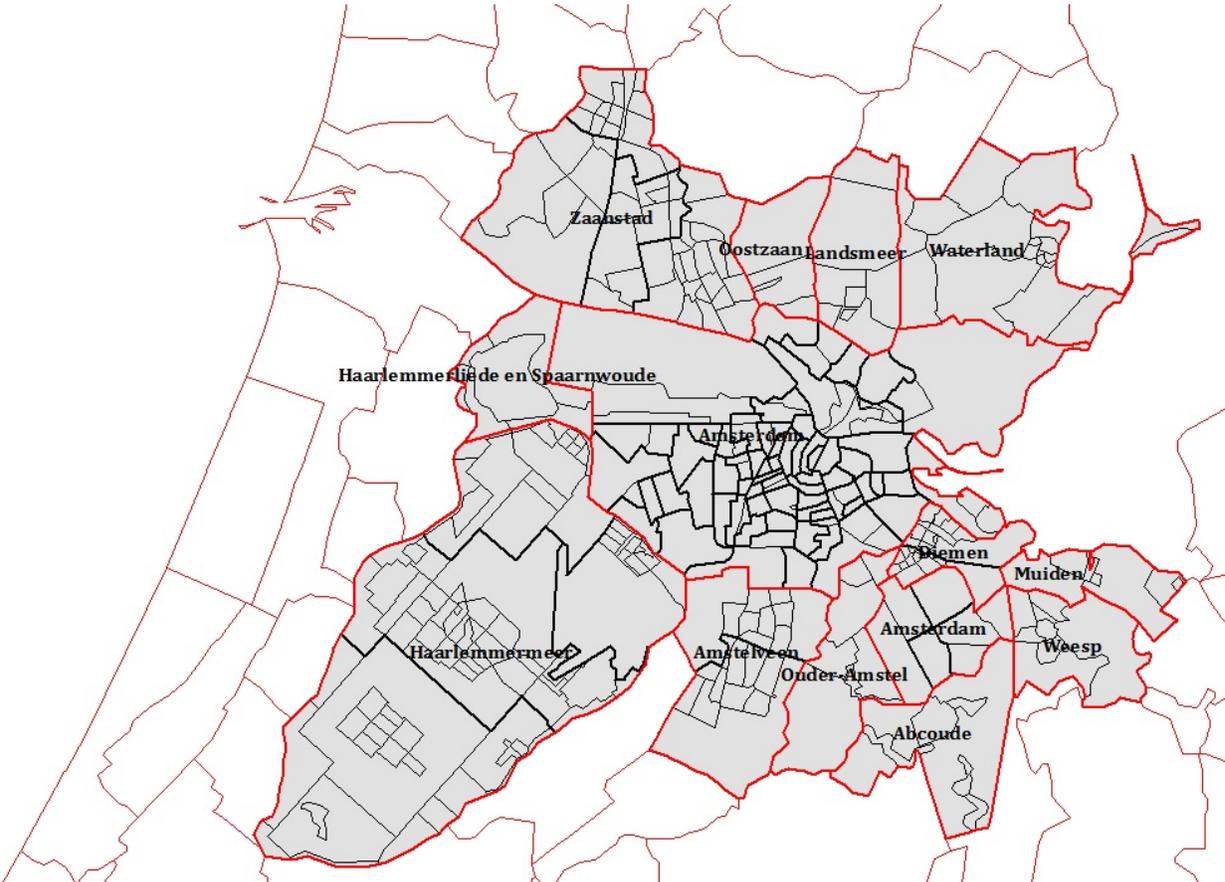


Figure A.1. Amsterdam and surrounding municipalities

Note: The thick red lines show the boundaries of the municipalities within our study area. The thin black lines show the boundaries of the administrative neighborhoods within each municipality. The thick black lines show the boundaries of the combinations of neighborhoods that have been used as choice alternatives within our modelling framework.

APPENDIX B. DERIVATIONS FOR BAYER'S INSTRUMENTS

The derivation presented here are related to the discussion in Section 2.3 of Weijsschede-van der Straaten et al. (2017). To see how Bayer's instrument for the price works with a heterogeneous population, we start from (1) and write the individual choice probabilities as:

$$\pi_j^i = e^{\delta_j + \varphi_j^i} / \sum_k e^{\delta_k + \varphi_k^i}. \quad (\text{B.1})$$

Taking logs on both sides and rearranging gives:

$$\delta_j + \varphi_j^i = \ln \pi_j^i + \ln \sum_k e^{\delta_k + \varphi_k^i}. \quad (\text{B.2})$$

The second term on the right-hand side is the expected maximum utility of consumer i , which we will denote as EU^i . Taking averages over the population, we find:

$$\delta_j = \overline{\ln \pi_j^i} + \overline{EU^i}, \quad (\text{B.3})$$

where we have used the fact that the φ_j^i s are on average equal to zero and have used a bar to indicate averages. Ignoring for the moment the 'social interaction' effect, we specify average utility as:

$$\delta_j = \alpha \ln p_j + \bar{\gamma} x_j + \xi_j. \quad (\text{B.4})$$

Substitution in (B.3) and rearrangement gives:

$$\ln p_j = (\overline{EU^i} + \overline{\ln \pi_j^i} - \bar{\gamma} x_j - \xi_j) / \alpha. \quad (\text{B.5})$$

In the estimated model (B.5) holds. It is important to note that $\overline{EU^i}$ and $\overline{\ln \pi_j^i}$ are both functions of the utilities and therefore of the (logged) prices and the unobserved characteristics. However, changes in $\overline{EU^i}$ imply an equal shift of all logged prices which leaves all choice probabilities unchanged, while changes in $\overline{\ln \pi_j^i}$ are j -specific and can have an impact on price differences.

We note that, because of the alternative-specific constants, it must be true that the average choice probability of j equals the share of the housing stock in j :

$$\overline{\pi_j^i} = s_j. \quad (\text{B.6})$$

This equality is disturbed if all ξ s are set equal to zero. The counterfactual prices that are Bayer's price instrument are computed by adjusting the price until the market equilibrium condition (B.6) is again satisfied. We write the counterfactual prices as p_j^{cf} and write them similar to (B.5) as:

$$\ln p_j^{cf} = (\overline{EU^{i^{cf}}} + \overline{\ln \pi_j^{i^{cf}}} - \bar{\gamma} x_j) / \alpha. \quad (\text{B.7})$$

Note that in (B.7) the unobserved characteristic does not occur because it has been set equal to 0, and that $\overline{EU^{i^{cf}}}$ and $\overline{\ln \pi_j^{i^{cf}}}$ have been given a superfix cf because their values will change when the unobserved characteristics are set equal to zero and prices are adjusted so as to re-establish a market equilibrium.

APPENDIX C. THE CANDIDATE INSTRUMENTS AND THEIR SELECTION

In this appendix we provide further information about the sets of candidate instruments and the outcomes of Belloni et al.'s (2012) selection procedure. Full results of first stage regressions are available from the authors upon request.

Squares and cross products

The set consists of the 9 squared exogenous variables and their 36 cross products. Selected instruments for the log price were the squares of variables 1 and 2, and the cross product of variables 2 and 9 and that of variables 6 and 9. For the share of high income households, only one candidate was selected: the cross product of variables 4 and 5.

Cubes and third order cross products

The set consists of the 9 cubed exogenous variables and their 84 third order cross products. For the price one candidate is selected: the cube of variable 2. For the share of high income households also only one candidate is selected: the cross product of exogenous variables 3, 4 and 6.

BLP-like instruments

The set consist of the 9 sums of the characteristics of neighborhoods at a distance of less than 5 km, less than 10 km and less than 15 km (27 candidate instruments), the 9 sums of the characteristics of neighborhoods at a distance of more than 5, 10 and 15 km (again 27 candidate instruments), and the 9 sums of the characteristics of the neighborhoods at a distance between 5 and 10, 10 and 15 and 5 and 15 instruments (27 'doughnut' instruments). Selected instruments for the price are the sum of characteristic 2 over neighborhoods at a distance of less than 5 km, the sum of characteristic 2 over neighborhoods at a distance of more than 5 km, and the sum of characteristic 1 over neighborhoods at a distance of more than 15 km. Selected instruments for the share of high income households are the sum of characteristic 7 over neighborhoods at a distance of less than 10 km and the sum of characteristic 7 over neighborhoods at a distance between 5 and 10 km.

Rank instruments

The instruments are the rankings of the neighborhoods based on each of the 9 characteristics. For the logged price the rank of characteristics 1, 2 and 9 were chosen, for the share of high income households those of characteristics 5 and 7.

Distance to an ideal neighborhood and number of close competitors

The ideal neighborhood is defined as one in which all characteristics have the most preferred value that exists in the set of neighborhoods. Distance measures were based on normalized values of the characteristics. Candidate instruments were: (i) the Euclidean distance to the ideal, (ii) Manhattan distance, (iii) the maximum of the difference to the ideal value for the 9 characteristics. The close competitors were defined as the number of neighborhoods at a Euclidean distance in characteristics space of at most 1, 2 or 3. In total this gives us 6 instruments based on normalized characteristics. For the log price the Manhattan distance was selected, for the share of high income households the maximum of the differences and the number of neighborhoods at a Euclidean distance of less than 3.

Bayer's instruments

The computation of these instruments has been discussed in the main text. The price instrument was selected for the log price, the other instrument for the share of high income households. In contrast to all sets of candidate instruments discussed above the first stages are very strong.

Other issues

When we consider all instruments jointly only Bayer's price instrument is selected for the log price and Bayer's share instrument for the share of high income households.

When we remove the two Bayer instruments from the set of candidates, the selected instruments for the price are the cross product of characteristics 2 and 9, the cross product of characteristics 1, 4 and 9, and the Manhattan distance to the ideal neighborhood. The selected instruments for the share of high income households are the cross product of characteristics 1, 4 and 6, the cross product of characteristics 1,5 and 8, and the sum of characteristic 7 for neighborhoods at a distance between 5 and 10 km.

APPENDIX D. REGRESSION RESULTS CONTINUED

Table D.1. First step results for renters: Deviations from mean indirect utility

Neighborhood characteristics	Household characteristics					
	High income households					
	Couple		With children (-18)		Age	
Conservation areas (km2)	-0.230 *	0.128	-0.141	0.143	0.022 ***	0.006
Conservation areas in surrounding neighborhoods	-0.168 ***	0.049	-0.199 ***	0.046	-0.009 ***	0.002
Distance to the nearest concentration of 100,000 jobs (km)	-0.019	0.019	0.052 ***	0.018	0.005 ***	0.001
Distance to the nearest intercity station (km)	0.008	0.027	-0.046 *	0.025	0.000	0.001
Nature (km2)	-0.074 **	0.037	-0.007	0.034	0.002	0.001
Nature in surrounding neighborhoods	-0.043 *	0.024	-0.098 ***	0.024	-0.004 ***	0.001
Water (km2)	0.231 ***	0.051	-0.123 ***	0.044	-0.007 ***	0.002
Water in surrounding neighborhoods	0.044	0.030	0.109 ***	0.027	0.009 ***	0.001
Age of the neighborhood	0.003 *	0.001	-0.004 **	0.002	-8.81E-05	6.34E-05
Log-likelihood	-133,928					
	Middle income households					
	Couple		With children (-18)		Age	
Conservation areas (km2)	-0.512 ***	0.097	-0.733 ***	0.124	0.012 ***	0.004
Conservation areas in surrounding neighborhoods	-0.267 ***	0.032	-0.125 ***	0.035	-0.001	0.001
Distance to the nearest concentration of 100,000 jobs (km)	-0.045 ***	0.013	0.020	0.014	0.003 ***	0.001
Distance to the nearest intercity station (km)	0.005	0.016	-0.055 ***	0.018	-0.001	0.001
Nature (km2)	-0.080 ***	0.024	-0.095 ***	0.027	0.002 *	0.001
Nature in surrounding neighborhoods	-0.090 ***	0.016	-0.109 ***	0.019	-0.001 *	0.001
Water (km2)	0.128 ***	0.030	0.025	0.033	-0.002 *	0.001
Water in surrounding neighborhoods	0.071 ***	0.019	0.071 ***	0.020	0.004 ***	0.001
Age of the neighborhood	0.002 *	0.001	-0.010 ***	0.001	-2.17E-04 ***	4.69E-05
Log-likelihood	-137,213					
	Low income households					
	Couple		With children (-18)		Age	
Conservation areas (km2)	-0.426 ***	0.081	-0.963 ***	0.095	0.011 ***	0.003
Conservation areas in surrounding neighborhoods	-0.128 ***	0.028	-0.168 ***	0.029	0.002 **	0.001
Distance to the nearest concentration of 100,000 jobs (km)	-0.031 ***	0.011	0.017	0.011	0.002 ***	0.000
Distance to the nearest intercity station (km)	0.003	0.014	-0.068 ***	0.015	0.000	0.001
Nature (km2)	-0.037 *	0.020	-0.108 ***	0.021	-0.003 ***	0.001
Nature in surrounding neighborhoods	-0.041 ***	0.014	-0.094 ***	0.015	-0.001 *	0.001
Water (km2)	0.050 *	0.026	0.064 **	0.027	0.001	0.001
Water in surrounding neighborhoods	0.036 **	0.015	0.072 ***	0.016	0.001	0.001
Age of the neighborhood	0.000	0.001	-0.010 ***	0.001	-1.79E-04 ***	3.39E-05
Log-likelihood	-135,997					

Note: Parameter estimates are used to calculate the deviations from the mean indirect utility with all variables normalized to have mean zero. 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.

APPENDIX E. DERIVATION OF THE MARGINAL WILLINGNESS-TO-PAY

Equation 1 can be rewritten almost as a hedonic price function allowing for heterogeneity in household preferences:

$$p_j^{g(i)} = \left(\frac{\beta^{g(i)}}{\alpha^{g(i)}}\right) \sigma_j + \left(\frac{\bar{\gamma}^{g(i)} + \gamma^i(z^i - \bar{z}^{g(i)})}{\alpha^{g(i)}}\right) x_j + \left(\frac{1}{\alpha^{g(i)}}\right) \xi_j^{g(i)} + \left(\frac{1}{\alpha^{g(i)}}\right) \epsilon_j^i - \left(\frac{1}{\alpha^{g(i)}}\right) u_j^i, \quad (\text{E.1})$$

where $p_j^{g(i)} = \ln(P_j^{g(i)})$. Using this equation, it is simple to compute the MWTP of each income group g and each type of household i within each group g for each of the neighborhood characteristics x_j :

$$\frac{\delta P_j^{g(i)}}{\delta x_j} = \left(\frac{\bar{\gamma}^{g(i)} + \gamma^i(z^i - \bar{z}^{g(i)})}{\alpha^{g(i)}}\right) P_j^{g(i)}. \quad (\text{E.2})$$

The household characteristics are constructed to have mean zero, $(z^i - \bar{z}^{g(i)})$. This simplifies the MWTP of the average household within each income group:

$$\frac{\delta P_j^{g(i)}}{\delta x_j} = \left(\frac{\bar{\gamma}^{g(i)}}{\alpha^{g(i)}}\right) P_j^{g(i)}. \quad (\text{E.3})$$