

Historic Amenities, Income and Sorting of Households

Hans R. A. Koster (SERC, Department of Spatial Economics, VU University, Amsterdam)
Jos Van Ommeren (Department of Spatial Economics, VU University, Amsterdam)
Piet Rietveld (Department of Spatial Economics, VU University, Amsterdam)

January 2013

This work is part of the research programme of the independent UK Spatial Economics Research Centre funded by a grant from the Economic and Social Research Council (ESRC), Department for Business, Innovation & Skills (BIS) and the Welsh Assembly Government. The support of the funders is acknowledged. The views expressed are those of the authors and do not represent the views of the funders.

© H. R. A. Koster, J. Van Ommeren and P. Rietveld, submitted December 2012

Historic Amenities, Income and Sorting of Households

Hans R. A. Koster*
Jos Van Ommeren**
Piet Rietveld**

January 2013

* SERC, Department of Spatial Economics, VU University, Amsterdam

** Department of Spatial Economics, VU University, Amsterdam

Acknowledgements

An earlier version of this paper was distributed under the title “Upscale Neighbourhoods: Historic Amenities, Income and Spatial Sorting of Households”. This paper benefited from a NICIS-KEI research grant. NVM (Dutch Association of Real Estate Agents), WDM and the Rijksdienst voor Cultureel Erfgoed are gratefully acknowledged for providing data. We thank Gabriel Ahlfeldt, Paul Koster, Stephen Ross, Jens Suedekum and the seminar audiences at the London School of Economics, the 2nd Workshop on Urban Economics (Barcelona), the 6th meeting of the Urban Economics Association (Miami) and Tinbergen Institute (Amsterdam) for useful comments.

Abstract

We test the impact of historic amenities on house prices and sorting of households within cities. Conservation area boundaries enable us to employ a semiparametric regression-discontinuity approach to measure the impact of historic amenities. The approach allows for household-specific preferences. Conditional on neighbour attributes, the price difference at the conservation boundary is about 3 percent. Internal historic amenities are also important, as listed houses are about 6 percent more expensive. It is shown that rich households sort themselves in conservation areas and in listed buildings, because they have a higher willingness to pay for historic amenities. The results contribute to an explanation for the substantial spatial income differences within cities.

JEL Classifications: R14, R21, R31, R38

Keywords: historic amenities, sorting, conservation areas, semiparametric regression-discontinuity design, hedonic price method

Historic Amenities, Income and Sorting of Households

By HANS R.A. KOSTER, JOS VAN OMMEREN AND PIET RIETVELD

I. Introduction

We observe huge differences in economic outcomes over space, also within cities. For example, households living in historic parts of Manhattan are about 75 percent wealthier than households living outside such areas, possibly due to the combination of historic amenities and sorting effects (Glaeser, 2011, pp. 150). Few empirical studies have investigated the effects of (historic) amenities on sorting of households. This may seem surprising, as income sorting may cause a plethora of problems related to social

segregation. In the classical monocentric city model of Alonso (1964), Mills (1967) and Muth (1967), it is assumed that households choose a residential location conditional on a workplace location. Rich households sort themselves in locations close to the central business district, depending on their preferences for house size and commuting costs (see e.g. Glaeser and Kahn, 2003; Glaeser et al., 2008). Wheaton (1977) shows that the income elasticity of housing demand and commuting time costs are of about equal value. Therefore, the monocentric city model fails to explain why income levels are generally increasing in distance to the city centre in the US, while this pattern is the reverse in most European cities (Mills and Lubuele, 1997; Anas et al., 1998; Brueckner et al., 1999). So, other factors than commuting costs, such as agglomeration economies, building attributes and external urban amenities, are likely to play a role in explaining easy-to-observe variation in household income over space.

In these theoretical studies it is typically ignored that many cities (in Europe and the US) were founded long ago (Anas et al., 1998). The building stock of cities can therefore be several centuries old. Glaeser and Gyourko (2005) and Glaeser et al. (2006) emphasise that economic theories that aim to explain the spatial structure of cities typically ignore their physical side. An exception is the paper by Brueckner et al. (1999), which develops a theory that explains the impact of exogenous amenities on urban spatial structure and, more specifically, on sorting of rich and poor within cities. In particular, they argue that historical amenities (e.g. old buildings) are likely to attract the rich to the city centres. Although Brueckner et al. (1999) provide some anecdotal support for their theory, a more rigorous empirical test is absent.

In this paper we test whether the historic building stock has an impact on house prices, and more specifically on sorting of rich and poor households. Brueckner et al. (1999) do not distinguish between external and internal historic amenities of residential buildings. In contrast, we distinguish between both types of amenities, because the effect of external amenities is particularly relevant for spatial policy. In essence, we adapt the hedonic price approach of Bajari and Kahn (2005), which essentially involves a two-step semiparametric estimation procedure. We deviate from this approach in two ways: first, we assume a stochastic bid price function, implying that the willingness to pay for a property is also a function of a privately observed housing attribute (e.g. number of friends in neighbourhood), which is not observed by the econometrician. This leads to imperfect sorting of households in the sense that we may observe households with identical observable characteristics and preferences in different houses. Second, we do not give a structural interpretation to the estimated implicit prices, because our attributes of interest

are dichotomous and we want to avoid strong functional form assumptions on the utility function. Our framework is also loosely related to the sorting framework of Bayer et al. (2007) and Ellickson (1981), but we use a hedonic price approach rather than a discrete-choice framework. In line with Bayer et al. (2007) we focus only on sorting on observable household characteristics.

More specifically, we employ a regression-discontinuity approach and use the boundary of *conservation areas* (or in the US: *historic districts*) to measure whether the willingness to pay (WTP) for *external* historic amenities differs between different income-groups. Historic amenities in these areas are generated by listed buildings, monuments, parks and the urban infrastructure from past times, but it is especially the combination of these features that generate amenities, which is called an *ensemble* effect (Brueckner et al., 1999; Evans, 1999; Ahlfeldt and Maennig, 2010; Lazrak et al., 2011). An important characteristic of conservation areas is that they are defined by *discrete* spatial changes in the level of historic amenities, so the boundaries of the conservation area are well-defined. These changes are discrete rather than continuous because city expansion is usually not gradual (so, not by house, but by district).¹ The amenities can be regarded exogenous for current residents and tend to have strong positive effects on house prices.² The pronounced positive effects are often very local (in one street or the neighbouring houses) and are strongly related to the view from one's house. One may interpret the effect of conservation as an external effect, under the identifying assumption that unobserved housing attributes are continuous at the conservation boundary. We also investigate whether the preference to reside in a listed building, which captures *internal* amenities, differs between households. When the rich have a stronger preference for historic (external and internal) amenities than the poor, we may expect to see sorting of rich households in conservation areas and in historic buildings.

This is not the first study that examines the house price effects of historic amenities. However, previous studies may suffer from omitted variable bias, limited geographical scope and do not control for household sorting. This likely leads to an overestimate of the

¹ This discreteness is particularly clear for urban areas created in the second half of the 19th century in continental Europe, medieval walls were removed (which strictly defined the cities' boundaries) and cities expanded very quickly during that period.

² Brueckner and Rosenthal (2009) show that in the US an older housing stock is an important reason why rich households migrate to the suburbs: older houses are often of lower quality and are located in city centres. In this study, which focuses on the Netherlands, we argue and show that a *much* older housing stock generates historic amenities and is a main factor that attracts rich households to the centre.

effect of historic amenities (see similarly Bayer et al., 2007).³ We employ a semiparametric regression-discontinuity design that enables us to reduce the omitted variable bias due to the lack of spatial controls, because we focus on areas very close to the conservation area boundary.⁴ In this way, we obtain the implicit prices of *external* historic amenities. The implicit prices for *internal* amenities are captured by a dummy indicating whether a building is listed. In this study, we allow the implicit prices to differ between households. Following Bayer et al. (2007), we also reduce the omitted variable bias by including a range of neighbour attributes and the distance to the boundary. This controls for the preference for unobservable neighbour attributes that vary continuously over space.

In a second stage, the willingness to pay is regressed on household characteristics to investigate the effect of household income, age and other household characteristics. Because conservation areas and listed buildings are dichotomous, we cannot point identify the willingness to pay using the estimated implicit prices. We therefore use interval regression and construct bounds on household-specific WTP-parameters.

The current estimation approach raises four important issues. First, our estimate of external historic amenities using the discontinuity approach likely provides an underestimate, because we ignore the (positive) effect of conservation areas on residences just outside the conservation area. So, we capture only (extremely) localised benefits of conservation to residents.⁵ Second, conservation areas usually imply restrictions regarding maintenance and limit possibilities for changes in the exterior of the house. This may negatively impact the house price and may lead to a downward bias of the effect of historic amenities. For this reason, we also investigate whether house prices change when an area becomes conserved. We did not detect a statistically significant impact of changes in conservation policies, suggesting that this bias is limited (see Appendix C). Third, one may argue that at the boundary, households not only sort on historic amenities but also on unobserved neighbour attributes. However, we emphasise that distance to boundary should capture the most important unobserved neighbour attributes. Moreover, unobserved neighbour attributes (e.g. race) are likely highly correlated with neighbour

³ See for some empirical studies, Asabere et al. (1989), Schaeffer and Millerick (1991), Clark and Herrin (1997), Coulson and Leichenko (2001), Leichenko et al. (2001), Navrud and Ready (2002), Coulson and Lahr (2005), Ahlfeldt and Maennig (2010), Lazrak et al. (2011).

⁴ One common approach in the hedonic price literature to capture spatial variation in house prices is to include property fixed effects. However, such an approach is not feasible in the current context because historical amenities hardly change over time, implying that the effect of conservation area is fully captured by the fixed effects, so that the effect of interest is not identified.

⁵ Clearly, there will also be other benefits (such as visiting tourists that value conservation areas) that are not included in our hedonic price approach based on the housing market.

income and educational degree, which we include in the analysis. Fourth, we cannot fully rule out the possibility that the conservation boundary effect is correlated with unobserved housing attributes. The effect at the boundary then not only captures external amenities, but also captures internal historic amenities related to the desirable attributes of the historic building occupied by the owner. However, we control for a wide range of housing attributes and show in Section IV that these housing attributes are not statistically significantly discontinuous at the conservation boundary.⁶ This increases the likelihood that unobserved housing attributes are also continuous at the conservation area boundary. We also include whether a building is listed, which should capture the most important internal historic amenities of a building. In any case, it is important to emphasise that this issue does not affect the conclusion that households sort on basis of historic amenities, whether these are external or internal.

We use a matched dataset based on high-quality data of house sales in and near urban conservation areas in the Netherlands, as well as data containing detailed household characteristics. So, information on housing transactions and household characteristics is combined. It is first shown that house prices and income are substantially higher in conserved areas. Unconditional on neighbour attributes, we then establish that house prices at the conservation boundary are about 10 percent higher. However, when we control for neighbour attributes (e.g. income, educational degree), the effect of conservation areas is much lower (about 3 percent). Again, we interpret this as a lower bound estimate of the effect of external historical amenities. In addition, we demonstrate that internal amenities are important. Households are willing to pay about 6 percent of the house price to reside in listed buildings. We allow the house price effect of the coefficients related to the conservation boundary and listed building dummy to differ among households. The differences in WTP-coefficients are then attributed to observed household characteristics. We find that a doubling of household income leads to a WTP for conservation areas that is 1.7 percentage points higher. We interpret this as the *total* effect of sorting on conservation areas. When we also allow for heterogeneity in the WTP for other housing attributes (e.g. richer households prefer larger houses and richer neighbours), doubling of income leads to an increase in the WTP for conservations areas of 0.8 percentage points. We interpret this as the *direct* effect of sorting on conservation areas. This effect may seem small, but note that this is still about 30 percent of the average effect of conservation areas on house prices, and, as we will show, the effect is likely an underestimate. Our evidence of the importance

⁶ Only the discontinuity in construction year is marginally significant at the 10 percent. We therefore will pay attention to this in the sensitivity analysis.

of internal and external historic amenities suggest that policies that aim to protect the historic building stock may have long-lasting effects on house prices and sorting of high income households in conserved historic districts. Additionally, we show that the WTP for listed buildings is also positively correlated with income.

This paper continues as follows. In Section II we discuss the estimation procedure. Section III discusses the data, which is followed by a graphical analysis in Section IV. Section V presents the results for the parametric specifications. Section VI considers the results of the semiparametric estimation procedure and Section VII concludes and derives some policy implications.

II. Empirical methodology

A. A stochastic bid price function approach

The approach of Bajari and Kahn (2005) consists of two stages. In the first stage, a nonparametric hedonic price function is estimated, which delivers implicit prices for housing attributes. In the second stage, utility parameters are recovered using the estimated implicit prices in the first stages and are regressed on household characteristics.⁷ We deviate from this framework in two respects. First, we will make different assumptions on the specification of the kernel, which has implications for the economic interpretation on the mechanisms behind household sorting. Second, we do not give a structural interpretation to the estimated WTP parameters because the attributes of interest are dichotomous and we aim to avoid strong functional form assumptions on the utility function.

To be more specific, in the first stage Bajari and Kahn (2005) use (nonparametric) local linear methods that rely on a kernel (see Fan and Gijbels, 1996). We emphasise that the choice of what variables to include in the kernel is directly related to fundamental assumptions about households sorting, because through the kernel it is specified which observations are similar to each other.⁸ It seems natural to assume that households with similar *housing outcomes* (in observables) have similar preferences. This is the assumption made by Bajari and Kahn (2005) when applying their approach to explain racial segregation in cities. Although this may be an appropriate assumption in some contexts (maybe in the context of racial segregation where extreme segregation is usually observed), one would prefer to allow for imperfect sorting, because it is seldom the case that identical

⁷ Bajari and Kahn (2005) refer to this as a three-stage approach.

⁸ In the hedonic price literature, local linear techniques are often applied, where parameters are estimated that depend on geographic location (also known as geographically weighted regression) or housing attributes (Bajari and Kahn, 2005; McMillen and Redfearn, 2010).

residences are occupied by households with identical (observable) characteristics. Another issue is that the latter assumption seems inefficient in the sense that many theoretical models indicate that identical households make different housing choices. For example, monocentric urban models with identical households imply that, in equilibrium, there are large household differences in housing *outcomes*: some households will choose to live in a small residence (close to the workplace), whereas other households will choose a large residence (far from the workplace) (Muth, 1969). So, although households make completely different choices, their utility functions may be the same.⁹

Alternatively, one may assume that households with similar *household characteristics* (in observables) have similar household utility functions, so the kernel depends on household characteristics. In the current paper, we apply this approach. This assumption is consistent with *imperfect sorting* because it is observed that households with different observable characteristics and identical preferences reside in different houses. This requires then a stochastic bid price function. More formally, assume a utility-maximising household h of type i that locates in a certain property x and receives utility u_{ix} from external historic amenities c_x , internal historic amenities ℓ_x , neighbourhood attributes n_x , housing attributes z_x , an housing attribute ξ_x which is commonly observed across households but unobserved by the econometrician and a composite good g_i .¹⁰ When an household of type i locates outside the city, it receives utility \bar{u}_i . Utility $u_{ix}(g_i, c_x, \ell_x, n_x, z_x, \xi_x)$ is maximised subject to the budget constraint $y_i = g_i + p_{ix}(c_x, \ell_x, n_x, z_x, \xi_x)$, where y_i denotes household income and p_{ix} denotes the price of for a property. The price of g_i is normalised to one. In equilibrium, a household of type i obtains the same utility everywhere, so \bar{u}_i , which defines its bid price function $p_{ix} = Y_i(c_x, \ell_x, n_x, z_x, \xi_x)$. Because income and reservation utility are the same for households of type i , we have suppressed the index h , y_i and \bar{u}_i in the above notation.

We allow for a stochastic instead of a deterministic bid price function, in the spirit of Ellickson (1981). The price p_{hix}^* a household h of type i is willing to pay is a function of the observed price p_{hix} and an unobserved random term, denoted by ϵ_{hix} . We interpret this random term as a *privately observed* housing attribute, so an attribute that is only observed by household h (e.g. a preference for the colour of the house, the presence of friends in the

⁹ Additionally, to follow the approach of Bajari and Kahn (2005) is quite computationally intensive, as a very large number of housing attributes have to be included in the kernel function to be estimated: in our dataset we have many housing and neighbourhood attributes.

¹⁰ We suppress a time component t in this exposition, but it will be included in the empirical section.

neighbourhood, the job location) and not by the econometrician (see discussions on the interpretation of ϵ_{hix} in Bayer and Timmins (2005) and Seim (2005)).¹¹ We then have:

$$(1) \quad \log p_{hix}^* = \log p_{ix} + \epsilon_{hix} = Y_i(c_x, \ell_x, n_x, z_x, \xi_x) + \epsilon_{hix},$$

where $Y_i(\cdot)$ is the observed bid price of a representative household i . To determine the probability that a household of type i occupies a property x , only the maximum bids from each type i are relevant. Let also assume that $Y_i(c_x, \ell_x, n_x, z_x, \xi_x)$ is linear in its parameters, so:

$$(2) \quad \log p_{ix}^* = \max_{h \in N_i} \log p_{hix}^* = \alpha_i c_x + \beta_i \ell_x + \gamma_i n_x + \delta_i z_x + \xi_x + \epsilon_{ix}^*,$$

where N_i is the total number of households of type i and $\epsilon_{ix}^* = \max_{h \in N_i} \epsilon_{hix}$. Because $\log p_{hix}^*$ is unobserved, we will only estimate the deterministic part of the bid price function. As we assume that $E[\epsilon_{ix}^* | c_x, \ell_x, n_x, z_x, \xi_x] = 0$, this does not affect the consistency of the parameters to be estimated.

It is insightful to assume that ϵ_{ix}^* is identically and independently distributed Weibull, because (2) leads then to a Logit model, although we do not make the assumption to estimate the parameters of the model.¹² The probability Π_{ix} that a household of type i occupies a property at x is then:

$$(3) \quad \Pi_{ix} = \frac{e^{\alpha_i c_x + \beta_i \ell_x + \gamma_i n_x + \delta_i z_x + \xi_x}}{\sum_{j=1}^J e^{\alpha_j c_x + \beta_j \ell_x + \gamma_j n_x + \delta_j z_x + \xi_x}},$$

where $j = 1, \dots, J$. It is then easy to observe that conditional on the presence of historic amenities c_x , the probability that a household of type i locates in x is a positive function of preferences α_i . Ellickson (1981) essentially uses (3) to identify the parameters of interest using a discrete choice framework. Again, we will estimate (2) directly, without making assumptions on the distribution of ϵ_{ix}^* .

So, one attractive feature of our framework is that it allows for *imperfect* sorting, which is a common feature in the housing market: two identical houses are often occupied by different types of households. We stressed earlier that the assumption of imperfect sorting is fundamentally different from Bajari and Kahn (2005). They estimate a nonparametric function where the preferences are estimated on basis of housing attributes, so $Y_x(\cdot) = \alpha_x c_x + \beta_x \ell_x + \gamma_x n_x + \delta_x z_x + \xi_x$. If households do not sort on unobservable characteristics and there is imperfect sorting in the sense that households with different

¹¹ Note that because ϵ_{hix} is a private housing attribute, it will not capitalise in the price paid p_{ix} by a household, although it is part of the willingness of a household to pay for a certain house.

¹² This distributional assumption seems reasonable (and not as arbitrary as in the standard conditional Logit setting), because when ϵ_{hix} is normal, lognormal, exponential or logistic distributed, ϵ_{ix}^* will be Weibull distributed (Gumbel, 1958).

observable housing characteristics occupy identical houses and assume that (3) holds, then $\hat{\alpha}_x = \sum_j \Pi_{jx} \alpha_j$. It is then easy to observe that the approach proposed by Bajari and Kahn (2005) leads to an underestimate of preference heterogeneity. In the case of perfect sorting (so that $\Pi_{ix} = 1$ for each property x), our estimation approach and the approach of Bajari and Kahn (2005) are asymptotically equivalent. In line with Bayer et al. (2007), we ignore unobservable household characteristics in the current setting but assume a private housing attribute ϵ_{ix}^* instead. If households mainly sort on unobservable household characteristics our approach may be inefficient and incorrect, whereas the approach by Bajari and Kahn (2005) may deliver correct estimates of preference heterogeneity.

As emphasised earlier, our estimates are interpreted as reduced form estimates, so in contrast to Bajari and Kahn (2005) we do not give a structural interpretation to the parameters α_i and β_i .¹³ This is mainly because c_x and ℓ_x are dichotomous housing attributes, so we cannot point identify the structural willingness to pay. Bajari and Kahn (2005) show that for dichotomous attributes, utility maximising implies a threshold decision rule: if the attribute is chosen, the willingness to pay is equal to or higher than the estimated implicit price, whereas if the attribute is not chosen, the willingness to pay is equal to or lower than the estimated implicit price. We come back to this issue in Section II.D. A second reason to avoid a structural interpretation of the estimated parameters is that strong functional form assumptions on the utility function are needed.¹⁴

B. Parametric estimation methodology

Before we turn to the two-stage nonparametric estimation procedure, we initially ignore preference heterogeneity, so we assume that $\alpha_i = \alpha$, $\beta_i = \beta$, $\gamma_i = \gamma$, $\delta_i = \delta$. To investigate the impact of historic amenities, one may regress floor space price on a dummy indicating whether a house is in a conservation area and whether the house is listed, while controlling for housing attributes, such as number of rooms and construction year, and neighbourhood attributes, such as the distance to the city centre, railway stations and highways. However,

¹³ Although we will not identify structural parameters, we will refer to WTP-parameters as preferences, because WTP-parameters are directly related to preferences of households (see also the next footnote).

¹⁴ If c_x would be a continuous variable, it may be shown that $\frac{\partial u_{ix}/\partial c_x}{\partial u_{ix}/\partial g_i} = \frac{\partial p_{ix}}{\partial c_{ix}}$, so the marginal rate of substitution between external historic amenities and the composite good is equal to the derivative of the house price with respect to c_{ix} . When one is willing to assume a linear utility function: $u_{ix} = A_i c_x + B_i \ell_x + \Gamma_i n_x + \Delta_i z_x + \Xi_x + g_i$, where $A_i, B_i, \Gamma_i, \Delta_i$ and Ξ_x are structural parameters of the utility function. It is then straightforward to see that $\hat{\alpha}_i = A_i/p_{ix}$, so the implicit parameter for external historic amenities is directly related to the structural willingness to pay A_i , which may be recovered from the data.

it is very likely that not all control variables, in particular relevant neighbourhood attributes, are observed and, hence, the estimates may suffer from omitted variable bias (Black, 1999).¹⁵ To address the problem of omitted neighbourhood variables, we employ a spatial regression-discontinuity approach. This implies that only observations that are close to a boundary that represents an *exogenous* difference in amenities are included (see similarly Black, 1999; Bayer et al., 2007). We then regress the logarithm of residential floor space price p (per square meter) of property x in year t on a dummy c_{xt} indicating whether the house is, or will be, part of a conservation area. To control for the innate beauty of historic buildings, we include whether a building is listed, denoted by ℓ_{xt} . This captures the implicit price of internal historic amenities. We include a number of control variables z_{xt} , including structural attributes (e.g. house size) and neighbourhood variables. To control for the preferences for pleasant neighbours, we include neighbour attributes n_{xt} . Specifically, we include the average neighbourhood income and the share of people with at least a bachelor's degree. We also include the distance to the boundary, which captures among other things (unobserved) neighbour attributes that vary continuously over space. Importantly, we include boundary area fixed effects θ_x , which control for unobserved neighbourhood attributes in the different conservation areas around the boundary (see similarly Bayer et al., 2007), and include transaction year dummies ϕ_t , which capture unobserved temporal effects:

$$(4) \quad \log p_{xt} = \alpha c_{xt} + \beta \ell_{xt} + \gamma n_{xt} + \delta z_{xt} + \theta_x + \phi_t + \xi_{xt},$$

where α , β , γ , and δ are parameters to be estimated and we interpret ξ_{xt} as an error term.¹⁶ We use a regression estimator that only takes into account observations that are close to the conservation boundary. This implies a weighted regression estimator where the weight $w_x = I(d_x < d_T)$, where $I(\cdot)$ is an indicator function that equals one when the condition is true (and zero otherwise), d_x is the distance to the conservation boundary and d_T is the threshold distance from the boundary (Hahn et al., 2001).¹⁷ Black (1999) points out that this methodology is equivalent to calculating differences in mean house prices on opposite sides of the boundary, *conditional* on control variables (see also Imbens and Lemieux,

¹⁵ Conservation areas may be attractive for unobserved reasons. For example, these areas are often near historical, and therefore current, employment centres.

¹⁶ There are other interpretations possible of ξ_{xt} . For example, Bajari and Benkard (2005) and Bajari and Kahn (2005) interpret this as an unobserved attribute. It is unlikely that in our case ξ_{xt} represents an unobserved attribute, because we include boundary area fixed effects.

¹⁷ To correct for a potential bias that is larger when observations are further away from the boundary (for example, because of omitted unobserved factors or location-specific preferences), it is sometimes preferred that observations near the boundary get more weight. We therefore also experiment with a bisquare weighting function in the sensitivity analysis.

2008). The main advantage of this approach is that the method reduces omitted variable bias. The main disadvantage is that one estimates the effect at the boundary of the conservation area, whereas one may be more interested in the average effect over all locations within the conservation area.

C. Semiparametric estimation methodology – first stage

To investigate heterogeneity in preferences, we employ a two-stage semiparametric estimation approach. We start assuming that households only sort on historic amenities and do have identical preferences for other housing and neighbourhood attributes. These unconditional differences between estimates of implicit prices may be interpreted as the (net) *total* sorting effect caused by historic amenities. The first stage encompasses the estimation of a partially linear hedonic price function in which historic amenities, captured by c_{xt} and ℓ_{xt} , are nonparametrically related to the price for floor space, in line with equation (2). The price function then becomes:

$$(5) \quad \log p_{ixt} = Y_i(c_{xt}, \ell_{xt}) + \gamma n_{xt} + \delta z_{xt} + \theta_x + \phi_t + \xi_{xt},$$

where $Y_i(\cdot)$ is some function of being in a conservation area, a listed building, conditional on housing attributes, boundary area and year fixed effects.

However, it may be argued that households also sort on other housing, and importantly, neighbour attributes. We therefore also estimate a semiparametric hedonic price function where the implicit price of n_{xt} and z_{xt} may be different between household types as to allow for substitution between housing attributes and historic amenities. The differences in implicit prices may be interpreted as the *direct* sorting effect of historic amenities, as we condition on sorting on other housing attributes.¹⁸ The price function then becomes:

$$(6) \quad \log p_{ixt} = Y_i(c_{xt}, \ell_{xt}, n_{xt}, z_{xt}) + \theta_x + \phi_t + \xi_{xt}.$$

We estimate $Y_i(\cdot)$ by local linear regression techniques (Fan and Gijbels, 1996). In what follows, we outline the estimation approach for estimating (6), but the estimation procedure to estimate (5) is computationally equivalent. Locally weighted regression is the most common nonparametric approach to analyse spatial data, as it allows for a flexible functional form and interactions between the variables included in the nonparametric function to be estimated and household characteristics h_i (McMillen and Redfearn, 2010). Because we use local linear regression techniques, we may write $Y_i(\cdot) = \alpha_i c_{xt} + \beta_i \ell_{xt} + n_{xt} \gamma_i + z_{xt} \delta_i$. In contrast to (4), the coefficients α_i , β_i and γ_i are now *household-type-*

¹⁸ Because the price of the rental property is a function of numerous fixed effects, it would be infeasible to estimate a fully nonparametric hedonic price function (Bontemps et al., 2008; McMillen, 2010), so we assume that location fixed effects and time fixed effects are homogeneous across the population.

specific. Note that household belong to the same type if they have exactly the same (observable) characteristics, which implies that households have the same preferences. In our data, the number of unique types is (slightly) lower than the number of households. In Appendix A we outline the detailed estimation procedure to estimate equations (5) and (6). It is also explained how the vector of bandwidths λ , which determine the smoothness of $Y_i(\cdot)$, are determined.

D. Semiparametric estimation methodology – second stage

Once we have estimated $\hat{\alpha}_i$ and $\hat{\beta}_i$ for each household, we will relate the estimated implicit prices to the willingness to pay for households to investigate whether high income households are willing to pay more for historic amenities. We assume a (log)linear relationship between the willingness to pay and the household characteristics h_i to facilitate interpretation. In principle, we aim to estimate:

$$(7) \quad \alpha_i^* = \zeta h_i + \mu_i,$$

$$(8) \quad \beta_i^* = \eta h_i + v_i,$$

where α_i^* and β_i^* denote the (unobserved) household-type specific willingness to pay for conservation areas and listed building, ζ and η are parameters to be estimated and μ_i and v_i denote household-type-specific error terms.

Bajari and Benkard (2005) and Bajari and Kahn (2005) argue that the willingness to pay of dichotomous attributes is not point-identified in the first stage: households cannot optimise the amount of conservation area or listed buildings consumed. Bajari and Benkard (2005) and Bajari and Kahn (2005) show that utility maximisation implies a threshold rule: when a household resides in a conservation area or listed building, we can infer that i 's WTP-parameter exceeds the implicit price for this characteristic, so $\alpha_i^* \geq \hat{\alpha}_i$ and $\beta_i^* \geq \hat{\beta}_i$, where α_i^* and β_i^* denote the (unobserved) willingness to pay. If a household chooses to live outside a conservation area or listed building, it holds that $\alpha_i^* \leq \hat{\alpha}_i$ and $\beta_i^* \leq \hat{\beta}_i$. A standard ordinary least squares (OLS) regression will yield inconsistent estimates of ζ and η , but with relatively low standard errors. More specifically, ordinary least squares typically provide underestimates of ζ and η because it understates the true heterogeneity in WTP-parameters (as it is assumed that $\alpha_i^* = \hat{\alpha}_i$). Bajari and Kahn (2005) assume that the second stage error term is normally distributed, so that they can use a Probit model where the coefficient related to the implicit prices is normalised to minus one. Given that $\mu_i \sim N(0, \sigma^2 I)$ and $v_i \sim N(0, \sigma^2 I)$, the Probit model will lead to consistent estimates of ζ and η , however, typically with rather large standard errors.

In our application, because of our relatively small dataset, we prefer to define upper and lower bounds for the WTP-parameters to identify the impact of housing characteristics. For example, for conservation areas, we define:

$$(9) \quad \underline{\alpha}_i = c_{xt} \hat{\alpha}_i + (1 - c_{xt})(\hat{\alpha}_i - \omega\sigma)$$

$$(10) \quad \bar{\alpha}_i = c_{xt}(\hat{\alpha}_i + \omega\sigma) + (1 - c_{xt})\hat{\alpha}_i$$

where $\underline{\alpha}_i$ and $\bar{\alpha}_i$ denote the lower and upper bound of the WTP-parameters, σ is the standard deviation of the standard normal distribution (see equation (11)) and ω is an exogenous parameter that defines the bounds of the distribution. When ω is large, the intervals are larger (and the standard errors are typically higher). We then estimate the effect of household characteristics (e.g. income) on the WTP for conservation areas by estimating ζ and σ , so we aim to estimate equation (7). The log-likelihood $\log \mathcal{L}(\cdot)$ is then given by:

$$(11) \quad \log \mathcal{L}(\zeta, \sigma) = \sum_{i=1}^I \log \left(\Phi \left(\frac{\bar{\alpha}_i(\omega, \sigma) - \zeta h_i}{\sigma} \right) - \Phi \left(\frac{\underline{\alpha}_i(\omega, \sigma) - \zeta h_i}{\sigma} \right) \right)$$

where $\Phi(\cdot)$ is the standard cumulative normal distribution with standard deviation σ .¹⁹ It is important to note that if we set $\omega = 0$, equation (11) boils down to an OLS regression, while when $\omega \rightarrow \infty$ equation (11) produces identical coefficients to the Probit model suggested by Bajari and Kahn (2005). This is illustrated in Appendix F. In the analysis, we will set $\omega = 2.5$, so that the WTP is maximally two and a half standard deviations below or above $\hat{\alpha}_i$, which seems reasonable. For listed buildings, we estimate equation (8) using the same procedure.

III. Data and context

A. Conservation areas

Conservation areas contain buildings and landscapes that are protected by the national government because of special architectural and historic interest. In the United States and United Kingdom there are already thousands of such areas (and the number is rising). Conservation area policies became an issue in the 1960's. In the United States, the government created a National Register of Historic Places. In United Kingdom, the term conservation area was introduced by the Civil Amenities Act in 1967. The Netherlands created its first conservation area in 1965. The protected surface area has increased by more than 300 percent since 1990 (see Figure 1). Nowadays, about 3.5 percent of total

¹⁹ Note that this is a special version of the log-likelihood function to estimate an interval regression, because the standard deviation σ also appears in $\bar{\alpha}_i$ and $\underline{\alpha}_i$.

residential area is designated as a conservation area, containing 4.8 percent of the residential properties.

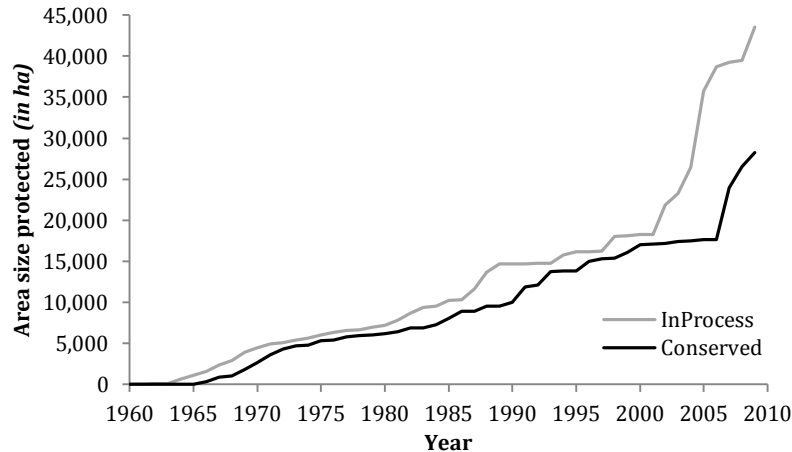


FIGURE 1 — TOTAL AREA SIZE DESIGNATED AS CONSERVATION AREA IN THE NETHERLANDS

In the Netherlands, designation of conservation area is the responsibility of the national government. The procedure for an area to be converted into a conservation area (in Dutch: *beschermd stadsgezicht*) is prepared by the municipality and the *Rijksdienst voor Cultureel Erfgoed* (RCE) (Department of Cultural Heritage). The RCE carefully determines the boundaries of conservation areas based on the presence of historic buildings, parks and urban infrastructure. So, conservation areas are almost always not just a concentration of listed buildings, but a combination of monuments, parks and urban infrastructure from past times, generating a positive ensemble effect (Ahlfeldt and Maennig, 2010; Lazrak et al., 2011). During the designation process, the RCE provides a document wherein they determine and clarify the exact location of the (future) conservation boundary. In a formal sense the municipality and house owners cannot influence this process, so the boundaries are thought to be exogenously determined (from the house owner perspective), in contrast to the US where local lobby groups may have a large impact on the definition of the boundaries of historic districts.²⁰ It is also important to note that these boundaries do not coincide with school attendance boundaries (as the latter are not strictly defined in the Netherlands) and usually also do not coincide with neighbourhood boundaries (as defined by Statistics Netherlands).

²⁰ So, it is generally not possible for house owners to influence this process by lobbying. Still, when there is lobbying, it is very unlikely that future house *buyers* have influenced the conservation process in the area in which they will buy the house.

Given these conservation boundaries, a municipality is obliged to develop a zoning plan which stipulates where changes to the housing stock are allowed, in such a way that it fits in the current historical landscape. As it is common practice in the Netherlands that a permit is needed for exterior changes of a house, the additional restrictions for existing buildings due to conservation areas are often limited. However, for developers intending to demolish existing buildings and/or construct new buildings, the conservation area status usually implies more restrictions.²¹

The process of transformation from an area that is not conserved to a conserved area officially takes 16 months, but usually takes longer with an average of 3 years and 4 months. Before designation, a certain area may already be protected by the local government. This may explain, for example, why the historic city centre of Amsterdam with its (famous) canals only recently (in 1999) became designated as a conservation area.

In this study, we focus on (national) conservation areas in urban areas in the Netherlands. The average size of the selected conservation areas is 75.6 hectare (the median is 57). We select conservation areas in cities that have more than 50,000 inhabitants and conservation areas in smaller cities that have an urban function (e.g. within a rural area), so we exclude rural conservation areas from the analysis.²² Figure B1 in Appendix B represents a map of the selected conservation areas. In the largest cities, such as Amsterdam, The Hague and Utrecht, historic city centres are protected. The city centre of Rotterdam was almost completely bombed in 1940, so only areas just outside the bombed area are conserved (see Koster et al., 2012). Figure B2 uses Rotterdam as an example and highlights that conservation areas are in some cases only one street, but may also comprise larger neighbourhoods.

B. Datasets

The preferred dataset would contain information on house prices and detailed characteristics (e.g. income) of the buyer of the house. We do not have such data, but we have two separate datasets, one on house prices and another on household characteristics. Using a large number of matching criteria regarding information that is available in both dataset, we are able to match households and houses.

²¹ Owners of houses in conservation areas are restricted in making changes to the exterior, whereas owners of listed buildings are also restricted in making inside changes. A positive aspect for the owner of a listed building is that subsidies are available for maintenance. Thus, the net effect of listing on buildings is ambiguous.

²² As the density of properties is much higher in large cities, this selection does not influence the results in any way.

The data on house prices is obtained from NVM (Dutch Association of Real Estate Agents). It contains information on a large majority (about 75 percent) of (owner-occupied) house transactions between 2002 and 2009. We know the transaction price, the exact address, and a wide range of structural attributes such as inside floor space size (in square meters), number of rooms, parking, garden, and construction year.²³ We also know whether the house is a listed building. Construction year controls for a range of difficult to observe attributes of houses (e.g. building quality, size of windows, architectural style, etc.). It is important to control for housing quality to be sure that the conservation boundary does not merely represent a discrete change in unobserved housing attributes. We therefore also include the state of maintenance, whether a house is insulated and has central heating. In the analysis, we control for a range of locational attributes. In particular, we include dummy indicators whether a property is within 100 meters of a major road, a railline, open space or water and within 250 meters of commercial buildings or a railway station.²⁴ Using a database from the RCE, we know whether a house is in, or close to, a conservation area. Our definition of conservation area includes areas that will be (or are known to be) designated as conservation areas. Including these areas is interesting, because historical amenities are very constant over time and do not alter when the area becomes a conservation area (in addition, in Appendix C we do not find evidence that changes in national conservation policies impact house prices). So, in the remainder of the paper, areas that will be designated will also be marked as ‘conservation areas’.

The data on household characteristics is obtained from a nationwide survey of WDM, a marketing service provider, which is held between 2004 and 2009. It consists of data on about 15 percent of all Dutch households and contains a wide array of household characteristics, such as gross income, number of children and household size, and also some characteristics of the head of the household, such as age, gender and educational level. The dataset also provides information on the house such as tenure status (price or rent) and type, and house location on a postcode six-digit level (PC6).²⁵ Gross monthly income is an ordinal-polytomous variable with seven categories. The latter variable is

²³ We exclude transactions with prices that are above € 1 million or below € 25,000 and have a price per square meter which is above € 5,000 or below € 500. We furthermore leave out transactions that refer to properties that are larger than 250 square meters or smaller than 25 square meters. These selections comprise less than one percent of the data.

²⁴ Land use data are for the year 2006 and are not year-specific. To use these data is reasonable because the land use pattern has hardly changed between 2002 and 2009, especially in or close to conservation areas.

²⁵ A postcode area contains on average 20 households and is comparable to a census block in the United States.

treated as continuous in our analysis, by taking the average of each class. Note that this data includes both renters and house owners.²⁶ Using these data, we are able to control for the average income and share bachelor's degree or higher in the neighbourhood of each property (excluding household's own income). This is useful as we aim to capture the direct effect of historic amenities, and not the indirect effect which may arise as historic amenities attract higher income and well-educated households. The historic-amenities induced increase in the average income of the neighbourhood may have a price increasing effect when households prefer to live close to rich households. We also calculate the household size, age and share of owner-occupied houses in the neighbourhood of each property.²⁷

We match the two datasets in order to gather information on the buyers of houses and their preferences for housing attributes. We select housing transactions that are within 500 meters of a conservation area boundary. This results in 162,925 housing transactions from the NVM sample and 28,569 household observations from the WDM sample. Then, the datasets are matched on six criteria: tenure status, postcode location, transaction year, house price and type, and construction year. First, only home owners are selected from the WDM dataset. We select housing transactions that took place maximum two years before the survey date. These households have recently made a moving decision which should reveal their actual preferences for housing attributes. Households that have not moved recently may not consume their optimal bundle of attributes because high transaction costs prevent them from moving to an optimal house (Bajari and Kahn, 2005). Second, we match on location: the NVM data provides exact locations of houses. When a house is in the same PC6-area, we match the data. Third, we match on house price. The NVM data uses a continuous transaction price, while the price of the WDM is a self-reported ordinal-polytomous price (10 classes). Houses and households are matched when the transacted price is in the correct interval. Fourth, we match on the construction year and house type. Because of our large number of matching criteria, we expect to correctly match households and houses.

The result of this matching procedure is a dataset of 4,958 house transactions and information on buyers' characteristics within 500 meters of conservation boundaries (so $d_T = 0.5$). Some descriptive statistics are presented in Table 1. It is shown that the average values are very comparable between the matched and full sample, so the matched sample

²⁶ The share of house owners is about 40 percent in our sample. In the Netherlands, the large majority of rents are controlled and owners of rent-controlled apartments are not allowed to sell these apartments by terminating rent controls. So, the share of ownership is most likely not determined in a free market.

²⁷ The educational level, gender and age refer to the head of the household.

seems to be a random selection of the full sample. Also for the WDM sample, the average values are very similar between the matched and full sample, although the share of listed buildings and houses in conservation areas is somewhat lower. In Appendix B we present descriptives for all housing attributes and household characteristics for the matched sample.

TABLE 1 — DESCRIPTIVE STATISTICS OF HOUSING TRANSACTIONS AND HOUSEHOLD CHARACTERISTICS

	Matched Sample		Full Sample NVM		Full Sample WDM	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Price per m ² (€)	2,087.891	635.395	2,194.903	761.987		
Conservation area	0.215		0.311			
Listed building	0.008		0.023			
House size (m ²)	101.636	34.521	104.776	40.666		
Apartment	0.538		0.581			
Rooms	3.926	1.292	3.865	1.548		
Maintenance good	0.873		0.838			
Central heating	0.884		0.854			
Construction year <1945	0.564		0.594			
Number of observations	4,958		162,925			
Gross Monthly Income (€)	3,512.859	1,626.031			3,673.361	1,708.470
Education – Bachelors or higher	0.467				0.467	
Household size	2.297	1.170			2.528	1.233
Age (year)	43.265	14.905			45.886	14.934
Number of observations	4,958				28,569	

Note: The samples only include observations that are within 500 meters of a conservation area boundary. We only present the most important housing attributes and household characteristics. For a full listing of descriptives, we refer to Appendix B.

The average price per square meter is € 2,088. The average price is € 2,308 and € 2,027 in respectively conserved and not conserved areas. So, unconditional on any other attribute, the average price per square meter is about 14 percent higher in conservation areas. Although suggestive of a conservation effect, it may be that this difference is due to sorting on neighbour attributes or price-increasing housing and neighbourhood attributes that are common in conservation areas, so one may not interpret this difference as an (external) historical amenity effect.

Household characteristics are also quite similar for conserved and not conserved areas, except for income and educational level. The average monthly gross income for the households in our sample (corrected for inflation) in conservation areas is € 3,880 and for non-conserved areas it is € 3,412. The average household size is the same as the Dutch average, with 2.1 and 2.3 in respectively conserved and non-conserved areas. The average age of the head of the household is 45 and 43, which is also close to the Dutch average.

IV. Graphical Analysis

We commence our empirical analysis by graphing the data close to the conservation boundary. Similar to Bayer et al. (2007), we regress the variables of interest on 50 meter distance band dummies. We graph the trendline by regressing the variable of interest on distance to the (nearest) conservation boundary (where negative distances indicate houses that are in the not conserved part) and a dummy indicating whether the property is in a conservation area. Figure 2 presents the results.²⁸

At the boundary there is a clear price difference (top-left panel), which suggests that the amenity value in conservation areas is substantially higher. When we include many housing attributes (more details will be provided in the next section), the effect is almost the same (top-right panel). It appears that the price difference *at the boundary* is about 10 percent. Our result that controlling for a large range of housing attributes and boundary area and year fixed effects does not have a large impact on the price difference at the boundary makes it likely that we do identify the effect of *external* historic amenities rather than the effect of unobserved housing attributes. However, note that this effect may still be an overestimate as we do not yet control for the preference for neighbours.

Discontinuous control variables do not invalidate the discontinuity design, because we estimate the effect of locating in the conservation area *conditional* on covariates (Imbens and Lemieux, 2008). Nevertheless, it is important to emphasise that differences between housing attributes are rather small or even do not exist at the boundary. Examples for house size, the number of rooms, the construction year and the state of maintenance are given in Figure 2. The share of houses that is constructed before 1945 is generally higher for areas that are conserved, but the discontinuity is only significant at a 10 percent level. The share of well-maintained houses is on average slightly (but not statistically significant) higher in conserved areas in line with the observation that conservation policies imply maintenance obligations. For rooms and house size (in square meters) we also do not find a statistically significant discontinuity at the conservation boundary.²⁹ In Figure 3 we employ the same procedure and graph the distribution of income and educational level along the

²⁸ Note that the number of observations within conservations areas decreases in distance to the conservation area boundary, as conservation areas are often rather small. Some dots may therefore contain only a small number of observations.

²⁹ One may still argue that there may be a correlation between construction year, state of maintenance, house size and the conservation area dummy. However, we will show that the effect of conservation is almost unaffected by inclusion of these housing attributes. Furthermore, we will show that our results are very similar when we only focus on houses built before 1945, well-maintained houses or houses smaller than 100 square meters.

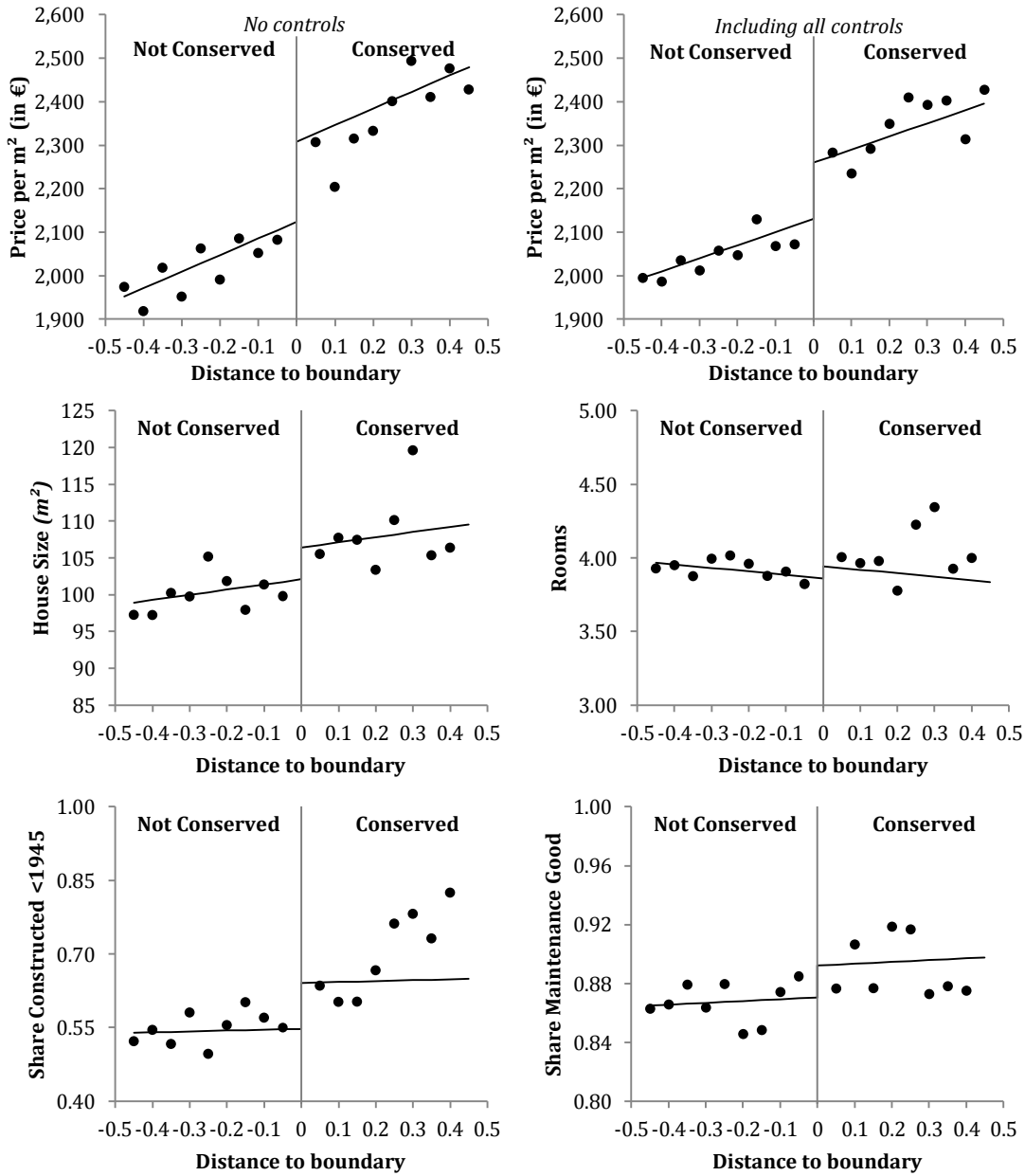


FIGURE 2 — DISCONTINUITIES ALONG THE CONSERVATION BOUNDARY

Notes: In the top-left panel, we regressed the price per square meter on 50 meter distance band dummies. In the top-right panel we also included all control variables (except distance to boundary). In the other panels, we regressed a housing attribute on 50 meter distance band dummies. So, a black dot represents a conditional average at a given distance from the conservation boundary. We also add a linear trendline that is estimated by a regression of the variable of interest on distance to boundary as well as a conservation area dummy.

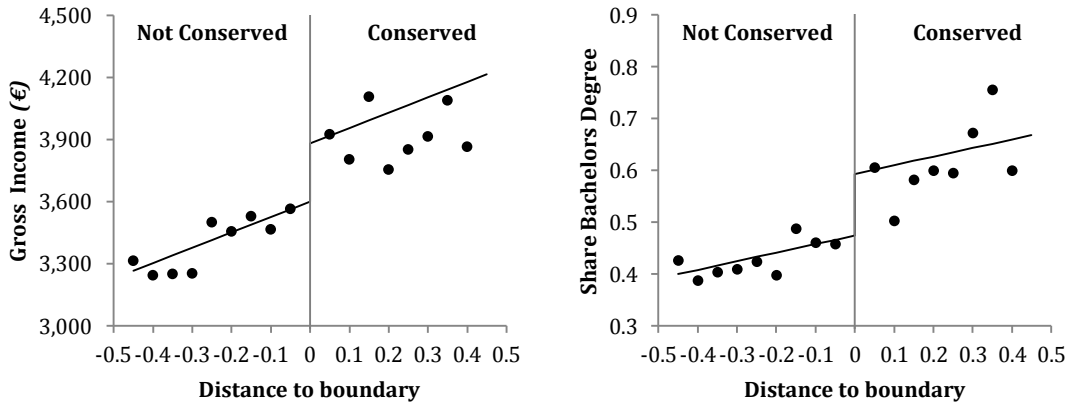


FIGURE 3 — DISCONTINUITIES OF HOUSEHOLD CHARACTERISTICS ALONG THE BOUNDARY

boundary for our sample. It is shown that income is on average 8 percent higher in conservation areas, and the share of people with at least a bachelor degree is 12 percentage points higher, which is not too surprising as educational level is strongly correlated with income and frequently also associated with demand for historical amenities (e.g. Glaeser, 1994; Glaeser et al., 2001; Shapiro, 2006; Florida et al., 2008).

It is therefore likely that households not only sort on basis of historic amenities at the conservation boundary but also on neighbour attributes. So we also investigate if there are discrete differences in *neighbour* attributes using the full WDM sample containing both renters and house owners. In Figure D1, Appendix D, we present the figures for relevant neighbour attributes. It is shown that especially neighbour income and neighbour educational degree are substantially higher in conservation areas and we observe discrete difference at the boundary of these variables. We also observe that the neighbour ownership share, neighbour household size and neighbour age are slightly lower in conservation areas. In the next section, we will investigate more thoroughly whether the price difference at the conservation boundary is attributable to the preferences for desirable characteristics of neighbours or historic amenities.³⁰

V. Results from parametric regressions

A. Regression results

In this section we will provide more evidence on the positive price effect of historic amenities. Table 2 presents the main results. The first parametric specification is the

³⁰ Note that neighbour variables are not perfectly collinear with historic amenities: the ensemble effect of historic amenities changes discontinuously at the boundary, but characteristics of neighbours also change within a conservation area enabling us to separately identify the effects.

standard approach to estimate the benefits of historic amenities, so we do not include neighbour variables n_{xt} and boundary area fixed effects θ_x . We include a conservation area dummy, whether a building is listed and a wide range of control variables. Then, the effect of conservation areas is 13.6 percent, so houses are substantially more expensive in conserved areas.³¹ Listed buildings also seem to be more expensive (14.7 percent). However, unobserved locational attributes may be correlated with the presence of conservation areas and listed buildings. Following Black (1999), in Specification (2) we therefore include boundary area fixed effects θ_x . It is shown that the effect of residing in a conservation area is then 11.0 percent. As emphasised above, it seems probable that we identify an ensemble effect of external historic amenities rather than the effect of unobserved housing attributes that are correlated with historic amenities, because we include a listed building dummy, boundary area fixed effects and a wide range of other housing attributes. The effect of residing in a listed building is 6.2 percent. Listed buildings are usually more aesthetically pleasing than non-listed buildings and receive some tax breaks for maintenance, which both explain the positive effect.

Discontinuities in income and educational level at the conservation boundary may lead to an overestimate of the effect of historic amenities because households may have a preference to locate near rich and well-educated households. We then estimate equation (3) and include neighbourhood income and the share of bachelor's degree in the neighbourhood as additional controls. To control for other neighbour attributes that are continuous at the boundary, we include a variable distance to the conservation boundary.³² It is shown in Specification (3) that the effect of residing in conservation areas is now substantially lower, but still 3.1 percent and highly statistically significant.³³ The effect of occupying a listed building is similar. In line with other studies, neighbourhood income is positively related to house prices. Our results suggest that doubling neighbourhood income increases houses prices by 10.8 percent. The educational level of neighbours also matters: one standard deviation increase in the share of people with a bachelor's degree increases house prices with 1.8 percent. However, neighbourhood incomes and educational level may

³¹ We calculate marginal effects as $e^{\hat{\alpha}} - 1$ and $e^{\hat{\beta}} - 1$.

³² We experiment with varying the geographical size of the 'neighbourhood' (e.g. 100 meters, 200 meters, etc.), but this has little impact on the results. Further, we also included the share of owners, average household size and average age of people in the neighbourhood, but this hardly impacts the results. These results can be received upon request.

³³ One may argue that it is preferable to cluster standard errors at the boundary area level. The effect of residing in a conservation area is still statistically significant at the 1 percent level. Residing in a listed building is then still statistically significant at the 5 percent level.

TABLE 2 — REGRESSION RESULTS ON THE IMPACT OF CONSERVATION AREAS
(Dependent variable: the logarithm of price per square meter)

	(1)	(2)	(3)
	Standard approach	Boundary approach	Boundary approach with neighbour attributes
Conservation area	0.128 (0.009)***	0.104 (0.008)***	0.031 (0.010)***
Listed building	0.137 (0.038)***	0.061 (0.027)**	0.059 (0.029)**
Distance to boundary (km)			0.143 (0.018)***
Neighbourhood income <100m (log)			0.156 (0.017)***
Share bachelor's degree <100m			0.084 (0.019)***
House size in m ² (log)	-0.224 (0.021)***	-0.262 (0.015)***	-0.301 (0.015)***
Rooms	0.013 (0.005)***	0.016 (0.003)***	0.017 (0.003)***
Terraced	0.033 (0.012)***	0.001 (0.010)	0.009 (0.010)
Semi-detached	0.231 (0.024)***	0.165 (0.019)***	0.143 (0.018)***
Detached	0.304 (0.056)***	0.289 (0.044)***	0.271 (0.042)***
Canal house	0.077 (0.059)	0.120 (0.045)***	0.129 (0.045)***
Manor	0.203 (0.020)***	0.160 (0.014)***	0.148 (0.014)***
Garden	0.111 (0.011)***	0.080 (0.008)***	0.074 (0.007)***
Balcony	0.038 (0.009)***	0.008 (0.006)	0.004 (0.006)
Roof terrace	0.024 (0.012)**	0.003 (0.009)	0.005 (0.008)
Parking	0.124 (0.013)***	0.119 (0.010)***	0.109 (0.010)***
Maintenance good	0.140 (0.012)***	0.118 (0.008)***	0.116 (0.008)***
Central heating	0.069 (0.012)***	0.066 (0.008)***	0.063 (0.008)***
Insulation	-0.003 (0.008)	-0.007 (0.006)	-0.003 (0.006)
Road <100m	-0.028 (0.007)***	-0.017 (0.005)***	-0.020 (0.005)***
Railway <100m	-0.005 (0.014)	-0.013 (0.010)	-0.015 (0.010)
Open space <100m	-0.005 (0.008)	-0.001 (0.006)	-0.002 (0.006)
Water <100m	0.036 (0.009)***	0.032 (0.007)***	0.025 (0.007)***
Commercial land use <250m	0.029 (0.007)***	-0.002 (0.006)	-0.002 (0.006)
Station <250m	-0.053 (0.021)**	-0.056 (0.016)***	-0.048 (0.015)***
Distance to centre (km)	0.022 (0.003)***	0.022 (0.007)***	0.030 (0.007)***
Construction year dummies (8)	Yes	Yes	Yes
Year fixed effects (6)	Yes	Yes	Yes
Boundary area fixed effects (115)	No	Yes	Yes
Number of observations	4,958	4,958	4,958
R ²	0.360	0.699	0.711

Notes: Distance to boundary is a variable ranging from -0.5 to 0.5, where negative distances indicate houses that are not in a conservation area. Share Bachelor's Degree <100m is the share of head of households that have a bachelor's degree or higher. The robust standard errors are between parentheses.

*** Significant at the 0.01 level

** Significant at the 0.05 level

* Significant at the 0.10 level

be positively related to omitted price-increasing housing and other neighbourhood attributes. Therefore, these estimates are not interpreted as causal effects (for example, neighbourhoods with on average high incomes often have lower crime rates and are also a proxy for the presence of good schools).

The control variables have in general the expected signs. The coefficient of distance to boundary is positive, implying that house prices increase when houses are closer to conservation areas.³⁴ This is not too surprising: households that live nearby a conservation area may visit the area at lower time costs and may reap some of the amenity benefits. However, distance to the boundary may also represent unobserved neighbour(hood) attributes. It is also shown that households' marginal WTP for house size is diminishing. Houses that are (semi-)detached, well-maintained, have a garden, a private parking space or a central heating are all more expensive, whereas houses near a major road or railway station are cheaper, *ceteris paribus*.

B. Sensitivity analysis

We now provide robustness checks of these results, where Specification (3) is used as a benchmark. We present the results in Table 3.

Omitted variables may bias our results. In the previous analysis we controlled for unobserved heterogeneity by including distance to the conservation boundary. Another way to control for unobserved spatial factors is to decrease the threshold distance d_T . Our threshold distance is arguably somewhat large. Researchers typically use a threshold distance of about 200 to 600 meters (Black, 1999; Bayer et al., 2007). When we reduce the threshold distance to 200 meters, it is still possible to estimate the effect of conservation areas with reasonable precision: the effect of conservation is very similar to Specification (3), as the effect is 4.1 percent (see Specification (4)). The effect of residing in a listed building is slightly higher and 7.8 percent.

In (5) we do not use a uniform weight function but a bisquare weight function, often applied in spatial applications, so observations closer to the boundary get more weight (McMillen, 2010).³⁵ It then appears that the effect of conservation areas is very similar to (3) and 3.3 percent. Specification (6) includes two distance-to-boundary variables instead

³⁴ We also note that houses in conservation areas that are further away from the conservation boundary are more expensive.

³⁵ Weights are then defined as: $w_x = (1 - (\delta_x/\delta_T)^2)^2 I(\delta_x < \delta_T)$.

of one, to allow for different slopes in conserved and non-conserved areas. The effect of conservation is equal to 3.5 percent.³⁶

Specification (7) considers the possibility that there is an error in the measurement of the conservation boundary. So, we exclude all properties within 25 meters on both sides of the conservation boundary. Then, the effect of conservation is 3.8 percent.

TABLE 3 — ROBUSTNESS ANALYSIS OF REGRESSION RESULTS
(Dependent variable: the logarithm of price per square meter)

	(4) <i>Threshold</i> $d_T = 0.2$	(5) <i>Bisquare</i> <i>weights</i>	(6) <i>Different</i> <i>slopes</i>	(7) <i>Measurement</i> <i>error</i>	(8) <i>Conservation</i> <i><2002</i>	(9) <i>Listed in</i> <i>vicinity</i>
Conservation area	0.040 (0.014) ***	0.032 (0.011) ***	0.034 (0.011) ***	0.037 (0.011) ***	0.031 (0.017) *	0.031 (0.010) ***
Listed building	0.075 (0.034) **	0.074 (0.031) ***	0.059 (0.029) ***	0.037 (0.028)	0.051 (0.037) ***	0.059 (0.029) **
Listed buildings <100m÷100						0.001 (0.056)
Control variables Included	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects (6)	Yes	Yes	Yes	Yes	Yes	Yes
Boundary area fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	2,741	4,958	4,958	4,675	1,841	4,958
R^2	0.725	0.722	0.720	0.714	0.679	0.711

Notes: See Table 2.

We examine whether focusing on areas that were already designated before 2002 changes the results. This is useful because areas that were designated more recently are potentially less attractive, for example because policies first focus on designation of the most attractive areas. Specification (8) highlights that the effect is very similar (3.1 percent). Our definition of conservation areas also includes areas that are in the process to become designated as a national conservation area. This may be not correct when areas that become conserved experience large changes in prices. In Appendix C, it is shown that we are not able to detect a statistically significant impact of recently introduced national conservation *policies* on house prices: the house price is not statistically different when a certain area becomes a formal conservation area. This suggests that (national) conservation policies simply capture the level of historic amenities present or that local policies already have protected certain areas before the official (national) designation.

Another concern may be that the discrete difference in price that we observe at the boundary represents a discrete difference in unobserved housing attributes, which may be correlated to observable housing attributes (e.g. state of maintenance, house size). When

³⁶ The coefficients of both distance variables are not statistically significantly different from each other, which justifies Specification (3).

we only select houses that are well-maintained, the effect is 3.2 percent. Limiting our sample to houses that are smaller than 100 square meters (to avoid the discontinuity in house size) results in an effect of 2.5 percent. When we only select houses that are constructed before World War II, the effect of conservation is 4.2 percent. So, this sensitivity analysis strongly suggests that potential discrete changes of housing attributes at the conservation boundary do not affect our results.

Finally, we include listed buildings in vicinity as an additional variable. It is shown that the effect is statistically insignificant (see (9)). This suggests that, conditional on residing in a conservation area, an additional listed building does not have an external effect.

VI. Results from semiparametric regressions

A. First-stage results

In this section, we investigate whether rich households have a stronger preference to locate in conservation areas. As explained in Section II, we employ a two-stage estimation procedure. In the first-stage, we estimate a semiparametric hedonic price regression to estimate implicit prices

We estimate four semiparametric specifications where we include all control variables, including neighbour, housing and location attributes (see Specification (3)). In Specification (10-1) we estimate equation (5) and initially assume that heterogeneity is only caused by income. In Specification (11-1) we include all household characteristics available to us in the kernel function. In Specification (12-1) we estimate equation (6), so we allow for sorting on neighbour and housing attributes. Again, we initially assume that income is the only determinant of preference heterogeneity. Specification (13-1) includes all household characteristics in the kernel function.

Table 4 presents averages and standard deviations of the coefficients that are estimated. The optimal bandwidth in Specification (10-1) is 0.8. The bandwidths for Specifications (11-1) are 1.0 for continuous household characteristics and 0.7 for categorical household characteristics. For Specifications (12-1) and (13-1) the bandwidths are similar. We refer to Appendix E for more details regarding the bandwidth selection and cross-validation scores.³⁷ It is shown that the average effects are similar to the estimates presented in Table

³⁷ There is no straightforward interpretation of the absolute values of the bandwidths, apart from the fact that these bandwidths capture the trade-off between fit and bias and indicate the similarity between individuals. Notice that when $\lambda_m \rightarrow \infty$ for continuous variables and $\lambda_m = 1$ for categorical variables, then there is no heterogeneity and the results will be identical to Specification (3).

2, although the effect of conservation is somewhat lower, suggesting that the parametric regressions lead to a small overestimate of the effect of conservation areas.

TABLE 4 — SEMIPARAMETRIC REGRESSION RESULTS
(Dependent variable: the logarithm of price per square meter)

	(10-1)		(11-1)		(12-1)		(13-1)	
	Total sorting effects		Total sorting effects		Direct sorting effects		Direct sorting effects	
	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev
Conservation area	0.021	0.009	0.017	0.028	0.024	0.005	0.024	0.021
Listed building	0.045	0.045	0.033	0.075	0.047	0.013	0.057	0.049
λ_m (continuous variables)	0.5		0.6		0.8		1.0	
λ_m (categorical variables)	n.a.		0.8		n.a.		0.7	
Control variables included	Yes		Yes		Yes		Yes	
Year fixed effects (6)	Yes		Yes		Yes		Yes	
Boundary area fixed effects (115)	Yes		Yes		Yes		Yes	
Number of observations	4,958		4,958		4,958		4,958	

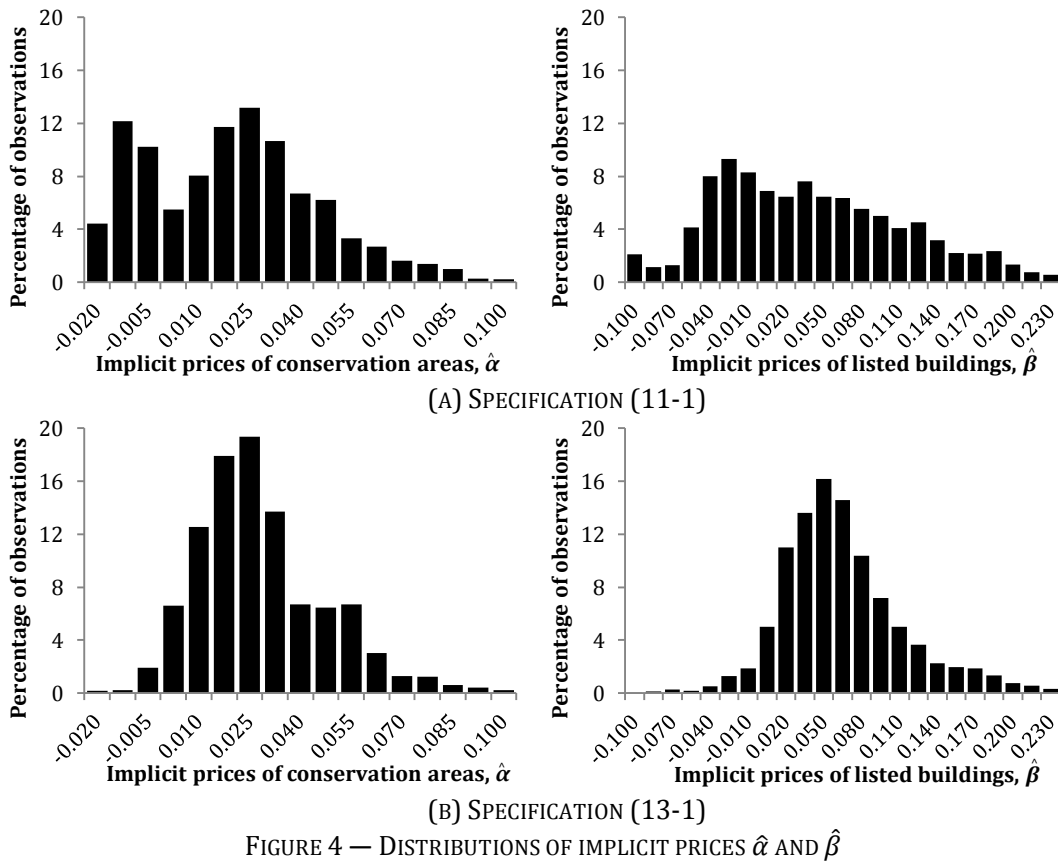


FIGURE 4 — DISTRIBUTIONS OF IMPLICIT PRICES $\hat{\alpha}$ AND $\hat{\beta}$

In Figure 4 the distributions of coefficients are shown. In Specification (11-1), about 30 percent of the households face a negative implicit price to locate in conservation areas. The implicit price for residing in a listed building is in 39 percent of the cases negative. In Specification (12-1), these percentages are lower, possibly because of higher bandwidths. About 6 and 7 percent of the households face a negative implicit price respectively to locate in conservation areas or to inhabit a listed building.

B. Second-stage results

In Table 5, we evaluate the second stage by estimating equation (7) and (8). We first evaluate sorting effects unconditional on sorting on other attributes. This may be interpreted as the (net) total sorting effect caused by historic amenities. Specification (10-2) only includes income in the kernel, while Specification (11-2) includes a wide range of other household characteristics. The dependent variables are latent WTP-parameters α_i^* and β_i^* . Throughout this subsection, we assume that the bounds parameter $\omega = 2.5$.

In the first stage Specification (10-1) we included income in the kernel but no other household characteristics. If this specification is valid, then, in the second stage it makes sense to only include income (and not any other household characteristics) as an explanatory variable of the WTP. It is shown that higher income households have a stronger preference for conservation areas than poor households: doubling household income leads to an increase in the preference for conservation areas of 1.5 percentage points. Doubling household income leads to a WTP for listed buildings that is 2.5 percentage points higher.

However, income may be strongly correlated to other household characteristics such as the level of education, age and the number of children. For this reason, in Specification (11-2) we include a large number of household characteristics (which are also included in the kernel in (11-1)). It is shown that the coefficient of income is similar to the one obtained with Specification (10-2) but slightly higher: doubling income leads to an increase in the WTP for conservation areas of 2.1 percentage points. It is important to note that income has a relatively strong impact on the preference for external historic amenities, compared to other household characteristics. When we regress the WTP-parameter of residing in a listed building $\hat{\beta}$ on household characteristics, the effect of income is positive and also slightly higher than in Specification (10-2).

TABLE 5 — SECOND STAGE INTERVAL REGRESSION RESULTS: EXPLAINING HETEROGENEITY, TOTAL SORTING EFFECTS
(Dependent variable: interval based on implicit prices for conservation areas $\hat{\alpha}$ and listed buildings $\hat{\beta}$)

	(10-2)		(11-2)	
	α_i^*	β_i^*	α_i^*	β_i^*
Income (<i>log</i>)	0.022 (0.000)***	0.036 (0.001)***	0.031 (0.002)***	0.041 (0.003)***
Education – secondary			0.004 (0.002)*	0.003 (0.003)
Education – vocational			-0.001 (0.002)	0.005 (0.003)*
Education – bachelors			0.004 (0.002)**	0.008 (0.003)**
Education – masters			0.013 (0.002)***	-0.011 (0.004)***
Dual earner household			-0.003 (0.002)*	0.019 (0.002)***
Household size			0.006 (0.001)***	-0.021 (0.001)***
Household – couple			-0.011 (0.002)***	0.017 (0.002)***
Household – couple w/ young kids			-0.001 (0.003)	-0.017 (0.004)***
Household – couple w/ older kids			-0.005 (0.003)	0.024 (0.004)***
Household – other			-0.004 (0.004)	0.055 (0.004)***
Male			-0.001 (0.001)	0.007 (0.002)***
Age 26-45			0.006 (0.002)**	-0.062 (0.003)***
Age 46-65			0.031 (0.002)***	-0.103 (0.003)***
Age >65			0.038 (0.003)***	-0.010 (0.005)**
Constant	-0.156 (0.003)***	-0.273 (0.007)***	-0.279 (0.014)***	-0.266 (0.021)***
σ	0.006 (0.000)***	0.021 (0.000)***	0.032 (0.001)***	0.050 (0.001)***
ω	2.5	2.5	2.5	2.5
Number of observations	4,958	4,958	4,958	4,958
Log-likelihood	-3551.172	-3484.467	-3464.834	-3497.794

Notes: See Table 2. The bootstrapped standard errors are between parentheses (1,000 replications). The reference category is a single female, younger than 26 with primary education. σ is the standard deviation of the standard normal distribution $\Phi(\cdot)$ and ω is the bounds parameter that, together with $\hat{\alpha}_i$, $\hat{\beta}_i$ and σ , defines $\underline{\alpha}_i$ and $\bar{\alpha}_i$ and $\underline{\beta}_i$ and $\bar{\beta}_i$.

Table 6 presents the results when we estimate equation (6) in the first stage, so we account for the possibility that households also sort on other housing and neighbour attributes. We interpret this as the direct sorting effect caused by historic amenities. Specification (12-2) only includes income in the kernel. It is shown that doubling of income leads to an increase in the WTP for conservation area of 0.5 percentage points. The effect is higher (0.8 percentage points) when we condition on other household characteristics (see Specification (13-2)). More specifically, when we compare a low income household (€ 1,500 per month) and a high-income household (€ 7,500 per month), the difference in WTP for residing in a conservation area is about 1.8 percent. Although this difference may seem quite small, note that it is still about 75 percent of the average effect of conservation areas on house prices (see Specification (12-1)). Furthermore, this is likely an underestimate because we set $\omega = 2.5$. In Appendix F, we therefore plot the effect of income against the choice of ω by re-estimating Specification (13-2) for different values of ω . It is shown that the Probit model becomes uninformative about the effect of income because of large

standard errors. It is therefore not statistically significantly different from the effect found in Specification (13-2).

TABLE 6 — SECOND STAGE INTERVAL REGRESSION RESULTS: EXPLAINING HETEROGENEITY, DIRECT SORTING EFFECTS
(Dependent variables: interval based on implicit prices for conservation areas $\hat{\alpha}$ and listed buildings $\hat{\beta}$)

	(12-2)		(13-2)	
	α_i^*	β_i^*	α_i^*	β_i^*
Income (<i>log</i>)	0.008 (0.000)***	0.013 (0.000)***	0.011 (0.001)***	0.027 (0.002)***
Education – secondary			0.001 (0.001)	-0.004 (0.002)**
Education – vocational			0.001 (0.001)	-0.007 (0.002)***
Education – bachelors			0.004 (0.001)***	-0.010 (0.002)***
Education – masters			0.013 (0.002)***	-0.033 (0.002)***
Dual earner household			-0.006 (0.001)***	0.020 (0.001)***
Household size			0.008 (0.001)***	-0.019 (0.001)***
Household – couple			-0.004 (0.001)***	0.025 (0.002)***
Household – couple w/ young kids			-0.001 (0.002)	0.003 (0.003)
Household – couple w/ older kids			0.002 (0.003)	0.033 (0.003)***
Household – other			-0.008 (0.002)***	0.036 (0.003)***
Male			0.003 (0.001)***	0.004 (0.001)***
Age 26-45			0.006 (0.002)***	-0.031 (0.002)***
Age 46-65			0.018 (0.002)***	-0.066 (0.002)***
Age >65			0.003 (0.002)	-0.010 (0.004)***
Constant	-0.045 (0.003)***	-0.073 (0.004)***	-0.107 (0.010)***	-0.131 (0.012)***
σ	0.006 (0.000)***	0.010 (0.000)***	0.023 (0.000)***	0.032 (0.001)***
ω	2.5	2.5	2.5	2.5
Number of observations	4,958	4,958	4,958	4,958
Log-likelihood	-3437.036	-3386.561	-3462.636	-3406.950

Notes: See Table 2. The bootstrapped standard errors are between parentheses (1,000 replications). The reference category is a single female, younger than 26 with primary education. σ is the standard deviation of the standard normal distribution $\Phi(\cdot)$ and ω is the bounds parameter that, together with σ , defines $\underline{\alpha}$ and $\bar{\alpha}$ and $\underline{\beta}$ and $\bar{\beta}$.

Income also has a statistically significant effect on the WTP for listed buildings. Doubling of income leads to an increase in the WTP for listed buildings of 0.9 or 1.9 percentage points in respectively Specifications (12-2) and (13-2). Note that education has a negative impact on the WTP for internal historic amenities but a positive impact on the WTP for external historic amenities. So, the results suggest that highly educated people tend to prefer to live in conservation areas, although they do not prefer to live in historic buildings. The same holds for older people.³⁸

³⁸ One may argue that multicollinearity may explain this result. However, we note that the variation inflation factors are all well below 5, suggesting that multicollinearity is not a major problem here.

VII. Conclusions

Few empirical studies have investigated the effects of amenities and historic building supply on sorting of households. This seems surprising given the large income differences within cities, which may lead to considerable problems that are associated with social segregation. In this paper we test whether external historical amenities may attract high income households to conservation areas. We assume a stochastic bid price function, leading to imperfect sorting of different types of households, dependent on the presence of external and internal historic amenities. Then, using a semiparametric regression discontinuity design, we investigate whether the WTP of different types of households for historic amenities differs between income groups, conditional on sorting effects. The test is performed in two steps.

First, using a discontinuity design that overcomes omitted variable bias and controls for neighbour attributes, we show that at the conservation boundary there is a price difference of about 3 percent. Because we do not find evidence that this effect is attributable to differences in housing quality at the boundary, we interpret this as a lower bound estimate of the effect of external historical amenities. So, under the identifying assumption that unobserved housing attributes are continuous at the conservation area boundary, we are able to evaluate the price effect of local external historic amenities. It is also shown that internal historic amenities have a positive price effect. More specifically, listed houses are about 6 percent more expensive.

Second, employing semiparametric regression techniques and using information on household characteristics, it is shown doubling of household income leads to a WTP for conservation areas that is 1.5 percentage points higher, consistent with the observation of sorting of high incomes along the conservation boundary. When we allow for sorting on neighbour and housing attributes, then it is shown that doubling of income leads to an increase in the WTP for conservation area of at least 0.8 percentage points. Although this difference seems quite small, note that it is still about 30 percent of the average effect of conservation areas on house prices and likely an underestimate. In a similar way, we find compelling evidence that rich households have a higher willingness to pay to reside in listed buildings.

From this paper, it may be concluded that the physical side of cities, and therefore history, leads to sorting of households over space. More specifically, policy makers should be aware of the fact that long-term national policies that stimulate preservation of historical buildings in specific cities may have large long-term spatial effects and may cause

problems related to social segregation, as high income households are attracted disproportionately by historic amenities.

References

- Ahlfeldt, G., Maennig, W. (2010). Substitutability and Complementarity of Urban Amenities: External Effects of Built Heritage in Berlin. *Real Estate Economics* 38(2): 285-323.
- Alonso, W. (1964). Location and Land Use: Toward a General Theory of Land Rent. *Cambridge MA: Harvard University Press*
- Anas, A., Arnott, R., Small, K.A. (1998). Urban Spatial Structure. *Journal of Economic Literature* 36(3): 1426-1464.
- Asabere, P.K., Hachey, G., Grubaugh, S. (1989). Architecture, Historic Zoning, and the Value of Homes. *Journal of Real Estate Finance and Economics* 2(3): 181-195.
- Bajari, P., Benkard, C.M. (2005). Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach. *Journal of Political Economy* 113(6): 1239-1276.
- Bajari, P., Kahn, M.E., (2005). Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities. *Journal of Business and Economic Statistics* 23(1): 20-35.
- Bayer, P.J., Ferreira, F., McMillan, R. (2007). A Unified Framework for Measuring Preferences for Schools and Neighbourhoods. *Journal of Political Economy* 115(4): 588-638.
- Bayer, P.J., Timmins, C. (2005). On the Equilibrium Properties of Locational Sorting Models. *Journal of Urban Economics* 57: 462-477.
- Bayer, P.J., Timmins, C. (2007). Estimating Equilibrium Models of Sorting Across Locations. *Economic Journal* 117: 353-374.
- Black, S.E. (1999). Do Better Schools Matter? Parental Evaluation of Elementary Education. *Quarterly Journal of Economics* 114: 577-599.
- Brueckner, J.K., Rosenthal, S.S. (2009). Gentrification and Neighbourhood Housing Cycles: Will America's Future Downtowns be Rich? *Review of Economics and Statistics* 91(4): 725-743.
- Brueckner, J.K., Thisse, J.F., Zenou, Y. (1999). Why is Downtown Paris so Rich and Detroit so Poor? An Amenity Based Explanation. *European Economic Review* 43: 91-107.
- Bontemps, C., Simoni, M. and Surry, Y. (2008). Semiparametric Hedonic Price Models: Assessing the Effects of Agricultural Nonpoint Source Pollution. *Journal of Applied Econometrics* 23: 825-842.
- Clark, D.E., Herrin, W.E. (1997). Historical Preservation Districts and Home Sale Prices: Evidence from the Sacramento Housing Market. *The Review of Regional Studies* 27: 29-48.
- Coulson, N.E., Lahr, M.L. (2005). Gracing the Land of Elvis and Beale Street: Historic Designation and Property Values in Memphis. *Real Estate Economics* 33(3): 487-507.
- Coulson, N.E., Leichenko, R.M. (2001). The Internal and External Impact of Historical Designation on Property Values. *Journal of Real Estate Finance and Economics* 23(1): 113-124.
- Ellickson, B. (1981). An Alternative Test of the Hedonic Theory of Housing Markets. *Journal of Urban Economics* 9(1): 56-79.

- Evans, A.W. (1999). The Land Market and Government Intervention. In: Cheshire, P., Mills, E.S. (eds.). *Handbook of Regional and Urban Economics 3, Applied Urban Economics: 1637-1669*.
- Fan, J., Gijbels I. (1996). Local Polynomial Modelling and Its Applications. In: *Monographs and Statistics and Applied Probability 66. Chapman and Hall, England*.
- Florida, R., Mellander, C., Stolarick, K. (2008). Inside the Black Box of Regional Development – Human Capital, the Creative Class and Tolerance. *Journal of Economic Geography 8: 615-649*.
- Glaeser, E.L. (2011). The Triumph of the City, How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier. *New York: Penguin Press*.
- Glaeser, E.L., Gyourko, J. (2005). Urban Decline and Durable Housing. *Journal of Political Economy 113(2): 345-375*.
- Glaeser, E.L., Gyourko, J., Saks, R.E. (2006). Urban Growth and Housing Supply. *Journal of Economic Geography 6(1): 71-89*.
- Glaeser, E.L., Kahn, M.E. (2003). Sprawl and Urban Growth. In: Henderson, J., Thisse, J.F. (eds.). *The Handbook of Urban Economics, vol. 4*.
- Glaeser, E.L., Kahn, M.E., Rappaport, J. (2008). Why do Poor Live in the Cities: the Role of Public Transportation. *Journal of Urban Economics 63(1): 1-24*.
- Glaeser, E.L., Kolko, J., Saiz, A. (2001). Consumer City. *Journal of Economic Geography 1: 27-50*.
- Gumbel, E.J. (1958). *Statistics of Extremes. New York: Columbia University Press*.
- Hahn, J., Todd, P., Van Der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica 69(1): 201-209*.
- Hurvich, C.M., Simonoff, J.S., Tsai, C.L. (1998). Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion. *Journal of the Royal Statistical Society B 60(2): 271-293*.
- Imbens, G.W., Lemieux, T. (2008). Regression Discontinuity Design: A Guide to Practice. *Journal of Econometrics 142: 615-635*.
- Ioannides, Y.M., Zabel, J.E. (2008). Interactions, Neighbourhood Selection and Housing Demand. *Journal of Urban Economics 63: 229-252*.
- Koster, H.R.A, Van Ommeren, J.N. and Rietveld, P. (2012). Bombs, Boundaries and Buildings: A Regression-Discontinuity Design to Measure Costs of Housing Supply Restrictions. *Regional Science and Urban Economics 42(4): 631-641*.
- Lazrak, F., Nijkamp, P., Rietveld, P., Rouwendal, J. (2011). Cultural Heritage: Hedonic Prices for Non-Market Values. *Mimeo, VU University Amsterdam*.
- Leichenko, R.M., Coulson, N.E., Listokin, D. (2001). Historic Preservation and Residential Property Values: An Analysis of Texas Cities. *Urban Studies 38(11): 1973-1987*.
- McMillen, D.P. (2010). Issues in Spatial Data Analysis. *Journal of Regional Science 50(1): 119-141*.
- McMillen, D.P., Redfean, C.L. (2010). Estimation and Hypothesis Testing for Nonparametric Hedonic House Price Functions. *Journal of Regional Science 50(3): 712-733*.
- Mills, E.S. (1967). An Aggregate Model of Resource Allocation in a Metropolitan Area. *American Economic Review 57: 197-210*.

- Mills, E.S., Lubuele, L.S. (1997). Inner Cities. *Journal of Economic Literature* 35: 727-756.
- Muth, R.F. (1969). Cities and Housing. *Chicago: University of Chicago Press*.
- Navrud, S., Ready, R.C. (2002). Valuing Cultural Heritage: Applying Environmental Valuation Techniques to Historic Buildings, Monument and Artifacts. *London: Edward Elgar*.
- Racine, J.S., Hart, J., Li, Q. (2006). Testing the Significance of Categorical Predictor Variables in Nonparametric Regression Models. *Econometric Reviews* 25(4): 523-544.
- Racine, J.S., Li, Q. (2004). Nonparametric Estimation of Regression Functions with both Categorical and Continuous Data. *Journal of Econometrics* 119(1): 99-130.
- Robinson, P.M. (1988). Root-N-Consistent Semi-Parametric Regression. *Econometrica* 57: 1403-1430.
- Sain, S.R., Baggerly, K.A., Scott, D.W. (1994). Cross-Validation of Multivariate Densities. *Journal of the American Statistical Association* 89(427): 807-817.
- Schaeffer, P.V., Millerick, C.A. (1991). The Impact of Historic District Designation on Property Values: An Empirical Study. *Economic Development Quarterly* 5(4): 301-312.
- Seim, K. (2006). An Empirical Model of Firm Entry with Endogenous Product-type Choices. *RAND Journal of Economics* 37(3): 619-640.
- Shapiro, J. (2006). Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital. *Review of Economics and Statistics* 88(2): 324-335.
- Van Ommeren, J.N., Wentink, D. (2011). The (Hidden) Cost of Employer Parking Policies. *International Economic Review, forthcoming*.
- Wheaton, W. (1977), Income and Urban Residence: An Analysis of Consumer Demand for Location. *American Economic Review* 67(4): 620-631

Appendix A. Estimation procedure

We use local linear regression techniques to estimate equations (5) and (6). In this subsection we focus on (6), but the procedure for (5) is similar. We estimate for each household-type i a weighted regression based on a multivariate kernel using household characteristics m :

$$(A1) \quad w_{ij} = \prod_{m=1}^M k_m(\lambda_m; h_{jm} - h_{im}),$$

where $m = 1, \dots, M$, and w_{ij} is the kernel weight of another household-type j in the local regression of i . $k_m(\cdot)$ is a kernel function of a chosen bandwidth λ_m and the difference between household characteristics of household-type j and i , so $h_{jm} - h_{im}$. Again, we do include household characteristics rather than housing attributes in the kernel. Most studies use the same kernel function $k_m(\cdot)$ for all variables. However, Racine et al. (2006) show that different kernel functions for continuous and categorical variables must be used. For continuous variables we employ the conventional Gaussian kernel function:

$$(A2) \quad k_{ijm} = \frac{1}{\lambda_m \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\tilde{h}_{jm} - \tilde{h}_{im}}{\lambda_m} \right)^2}.$$

Where k_{ijm} is the kernel weight of household-type j in the local regression of household-type i of characteristic m , \tilde{h}_{im} denote standardised household characteristics (with mean zero and unit standard deviation). For dichotomous and categorical variables we employ the following kernel function (see Racine and Li, 2004; Racine et al., 2006):

$$(A3) \quad k_{ijm} = \begin{cases} 1 & \text{if } h_{jm} = h_{im} \\ \lambda_m^{|h_{jm} - h_{im}|} & \text{if } h_{jm} \neq h_{im}. \end{cases}$$

The choice of the values of bandwidths λ_m is important (Sain et al., 1994; Hurvich et al., 1998). Lower bandwidths lead to a lower bias, but to higher variance of the estimator. Larger bandwidths may create a larger bias when the underlying function is nonlinear (Fan and Gijbels, 1996). We use a cross-validation procedure to determine the bandwidths (see Racine et al., 2006). A cross-validation score (\mathcal{CV}) is then defined as:

$$(A4) \quad \mathcal{CV}(\lambda_1, \dots, \lambda_M) = \sum_x (\log p_{ixt} - \log \hat{p}_{-ixt}(\lambda_1, \dots, \lambda_M))^2,$$

where $\log \hat{p}_{-ixt}(\lambda_1, \dots, \lambda_M)$ denotes the fitted value of $\log p_{ix}$ with the observation of i omitted from the calibration process. We choose the bandwidths $\lambda_1, \dots, \lambda_M$ that minimise the cross-validation score. To keep the computation process tractable, we use one bandwidth for continuous variables and another for categorical variables.

As equation (6) is partially linear, we employ the Robinson procedure (Robinson, 1988). First, we regress $\log p_{ixt}$, boundary and year dummies on c_{xt} , ℓ_{xt} , n_{xt} and z_{xt} nonparametrically. Then, we regress the residuals of $\log p_{ix}$ on the residuals of boundary and year dummies, which leads to \sqrt{N} -consistent estimates of θ_x and ϕ_t .³⁹ The last step is to regress $\log p_{ixt} - \hat{\theta}_x - \hat{\phi}_t$ nonparametrically on c_{xt} , ℓ_{xt} , n_{xt} and z_{xt} to get an estimate for $Y_i(\cdot)$.

³⁹ Robinson (1988) showed that the coefficients are estimated at parametric rates of convergence, despite the presence of a nonparametric part.

Appendix B. Context and descriptives



FIGURE B1— STUDY AREA

TABLE B1 — DESCRIPTIVE STATISTICS OF HOUSING TRANSACTIONS

	Full Sample		Sample, Conserved		Sample, Not Conserved	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Price per m ² (€)	2,087.891	635.395	2,308.398	680.757	2,027.423	608.667
Conservation area	0.215		1.000		0.000	
Conservation area – not designated	0.075		0.351		0.000	
Conservation area – completed	0.046		0.216		0.000	
Conservation area – in process	0.093		0.434		0.000	
Listed building	0.008		0.025		0.004	
Listed buildings <100m	1.081	5.314	4.310	10.658	0.196	1.098
Neighbourhood income <100m (€)	3,205.926	830.813	3,438.270	936.381	3,142.212	787.677
Share bachelor's degree <100m	0.389	0.216	0.500	0.214	0.359	0.207
Distance to boundary (km)	-0.163	0.220	0.163	0.117	-0.253	0.143
House size (m ²)	101.636	34.521	106.421	41.488	100.324	32.232
Apartment	0.538		0.598		0.522	
Terraced	0.408		0.351		0.424	
Semi-detached	0.048		0.044		0.049	
Detached	0.006		0.007		0.005	
Canal house	0.003		0.014		0.001	
Manor	0.066		0.112		0.053	
Rooms	3.926	1.292	3.944	1.547	3.921	1.213
Garden	0.508		0.483		0.515	
Balcony	0.482		0.516		0.473	
Roof terrace	0.096		0.108		0.093	
Parking	0.151		0.165		0.147	
Maintenance good	0.873		0.892		0.867	
Central heating	0.884		0.888		0.883	
Insulation	0.718		0.657		0.734	
Road <100m	0.473		0.480		0.471	
Railway <100m	0.075		0.046		0.083	
Open space <100m	0.250		0.305		0.235	
Water <100m	0.207		0.258		0.194	
Commercial land use <250m	0.572		0.687		0.540	
Station <250m	0.022		0.037		0.018	
Distance to centre (km)	1.632	1.069	1.400	0.846	1.696	1.114
Construction year <1906	0.040		0.074		0.031	
Construction year 1906-1930	0.285		0.262		0.292	
Construction year 1931-1944	0.239		0.305		0.221	
Construction year 1945-1959	0.093		0.064		0.101	
Construction year 1960-1970	0.062		0.015		0.075	
Construction year 1971-1980	0.047		0.053		0.045	
Construction year 1981-1990	0.075		0.088		0.071	
Construction year 1991-2000	0.103		0.089		0.107	
Construction year >2000	0.057		0.050		0.059	
Number of observations	4,958		1,067		3,891	

Note: Distance to boundary is a variable ranging from -0.5 to 0.5, where negative distances indicate houses that are not in a conservation area.

TABLE B2 — DESCRIPTIVE STATISTICS OF HOUSEHOLD CHARACTERISTICS

	Full Sample		Sample, Conserved		Sample, Not Conserved	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Gross Monthly Income (€)	3,512.859	1,626.031	3,880.628	1,843.942	3,412.008	1,546.019
Education – Primary	0.121		0.063		0.137	
Education – Secondary	0.198		0.184		0.202	
Education – Vocational	0.214		0.160		0.228	
Education – Bachelors	0.304		0.341		0.294	
Education – Masters	0.163		0.252		0.138	
Dual Earner Household	0.423		0.407		0.427	
Household Size	2.297	1.170	2.148	1.126	2.337	1.178
Household – Single	0.338		0.383		0.325	
Household – Couple	0.355		0.366		0.352	
Household – Couple w/ Young Kids	0.196		0.150		0.209	
Household – Couple w/ Older Kids	0.061		0.052		0.064	
Household – Other	0.050		0.049		0.050	
Male	0.645		0.609		0.655	
Age (year)	43.265	14.905	44.672	15.481	42.879	14.722
Number of Observations	4,958		1,067		3,891	

Note: The educational level, gender and age refer to the head of the household.

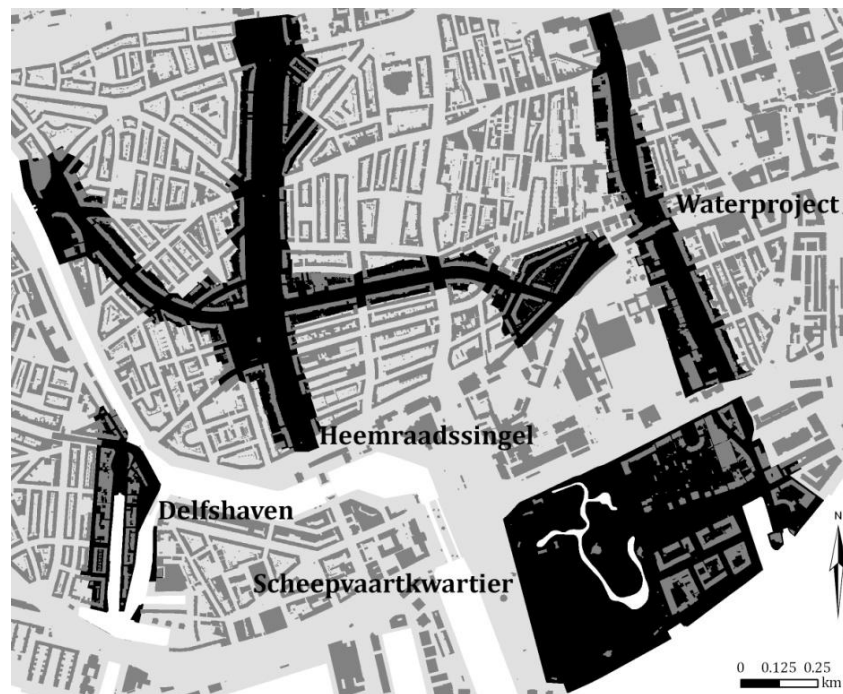


FIGURE B2— CONSERVATION AREAS IN ROTTERDAM

Appendix C. Do national conservation policies impact house prices?

Our definition of conservation areas includes areas that will become designated as a national conservation area. We may therefore evaluate price effects of national conservation policies (conditional on any local policy that is present; for example height restrictions that prevent new development). Regulations in the form of maintenance obligations and exterior restrictions may on the one hand induce positive externalities when neighbouring houses also have to maintain their houses, leading to a higher quality of the neighbourhood, while on the other hand they may imply an implicit tax for the house owner (Coulson and Lahr, 2005; Ahlfeldt and Maennig, 2010).

The ideal identification strategy would be to employ a repeat sales methodology that effectively compares the price of a house when it is conserved with the price when it is not yet conserved, conditional on the time trend. This essentially means that a residence fixed effect strategy is employed. As this will lead to an extremely small sample, we do not include residence fixed effects, but postcode six-digit fixed effects. These are very small areas, so we may assume that unobserved heterogeneity with respect to the environment does not play a role (see similarly Van Ommeren and Wentink, 2011).⁴⁰

TABLE C1 — REGRESSION RESULTS ON THE IMPACT OF A CHANGE IN CONSERVATION AREA STATUS
(Dependent variable: the logarithm of price per square meter)

	(C1)	(C2)	(C3)
Conservation Area – In Process	0.009 (0.028)	-0.014 (0.028)	0.003 (0.021)
Conservation Area – Completed	-0.036 (0.062)	-0.078 (0.060)	-0.025 (0.062)
Listed Building	0.024 (0.105)	0.115 (0.081)	0.078 (0.037)*
Housing Attributes (22)	Yes	Yes	Yes
Neighbourhood Attributes (9)	No	No	Yes
PC6-Area Fixed Effects (3,638/828)	Yes	Yes	No
Conservation Area Fixed Effects (75)	No	No	Yes
Year Fixed Effects (6)	Yes	Yes	Yes
Number of Observations	4,958	1,067	1,067
R^2	0.973	0.979	0.730

Notes: See Table 2. The reference category is ‘Conservation Area – Not Designated’. The standard errors are clustered at the conservation area level.

We present the results in Table C1. In Specification (C1) we include 3,638 PC6-area fixed effects. A distinction is made between areas that are not (yet) designated, that are in the process of becoming designated and areas where the conservation process has been completed. About 15 percent of the observations experience a change in the status of the conservation area. The results do not show a statistically significant different effect when an area becomes conserved. The sign of the coefficient after designation is even negative,

⁴⁰ The average distance to the centroid of a PC6-area is only 24 meters in our sample.

suggesting that the implicit regulation tax due to the national conservation policy offsets potential positive externalities of conservation to households. In Specification (C2) we only focus on observations that are either in a conservation area or will be in the future (leading to 828 PC6-area fixed effects). It is shown that the effect of becoming conserved is then somewhat stronger but still statistically insignificant. This is not too surprising, as we have only 1,067 observations, 828 fixed effects, and 30 control variables, implying few degrees of freedom. We therefore include conservation area fixed effects instead of PC6-area fixed effects in Specification (C3). The effect of becoming conserved is still (negative and) statistically insignificant.

So, we are not able to detect a statistically significant impact of conservation policies based on our data. One explanation for these results is that before designation, areas are already protected by the local government, so that little will change when national conservation policies are implemented. Another explanation is that regulatory constraints, such as height restrictions, zoning restrictions, preservation of open space and prevention of land use fragmentation are likely to increase developers' costs leading to lower land rents, but do not necessarily impact house prices (Ihlandfeldt, 2007; Glaeser and Ward, 2009).

Appendix D. Discontinuities of neighbour attributes

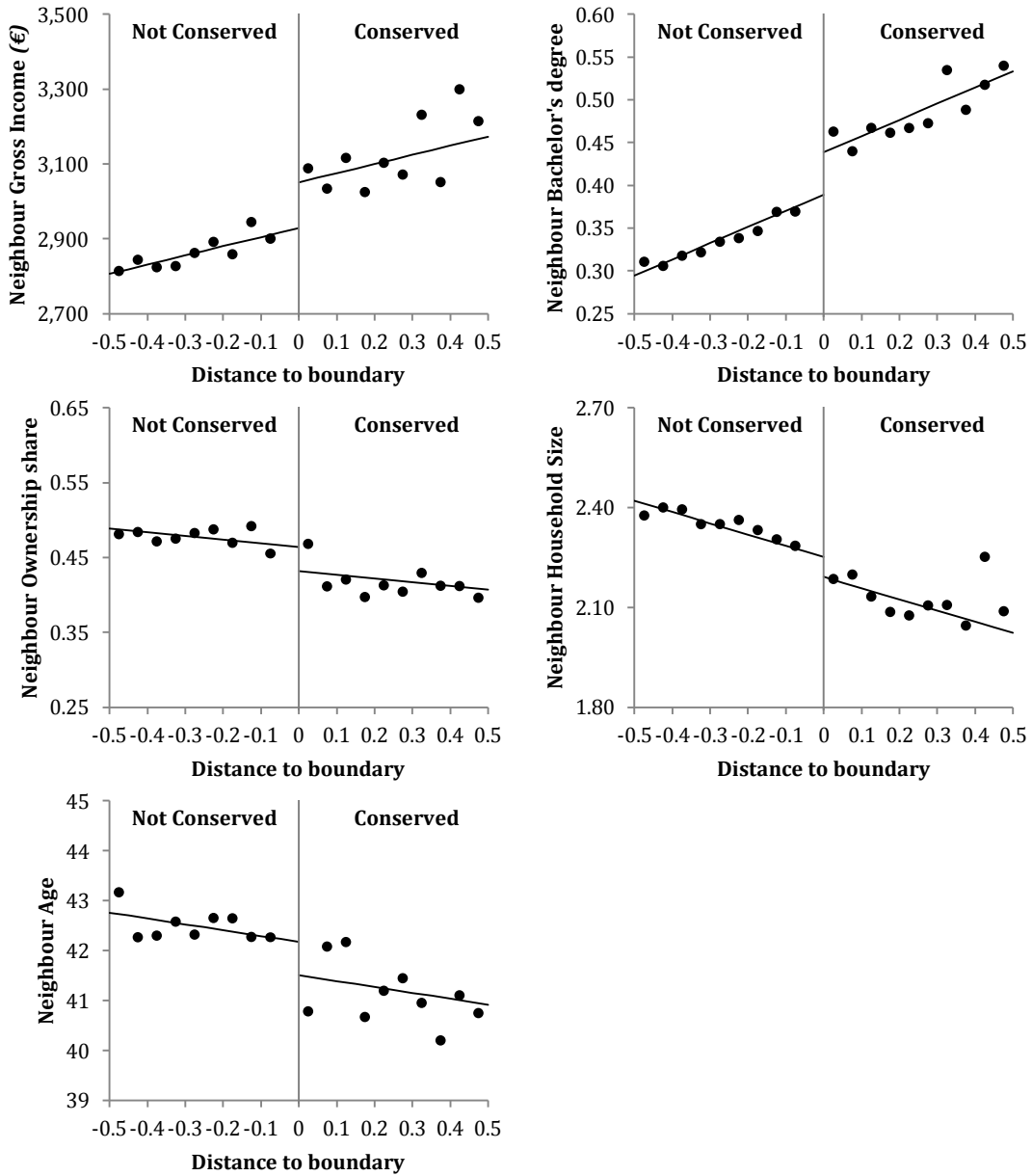


FIGURE D1 — DISCONTINUITIES OF NEIGHBOUR ATTRIBUTES ALONG THE CONSERVATION BOUNDARY

Notes: The graphs are based on the full dataset of WDM, including both house owners and renters. See also Figure 2.

Appendix E. Cross-validation scores and bandwidth selection

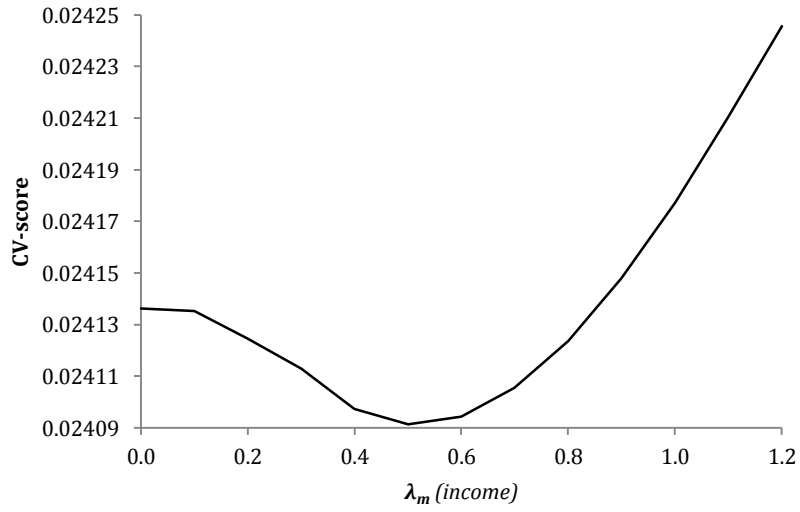


FIGURE E2 — BANDWIDTH SELECTION FOR SPECIFICATION (10)

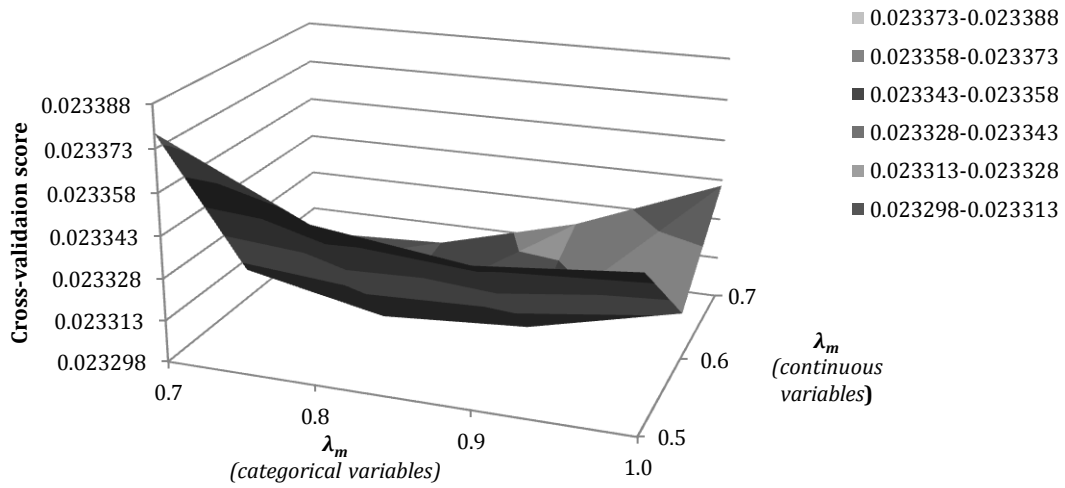


FIGURE E1 — BANDWIDTH SELECTION FOR SPECIFICATION (11)

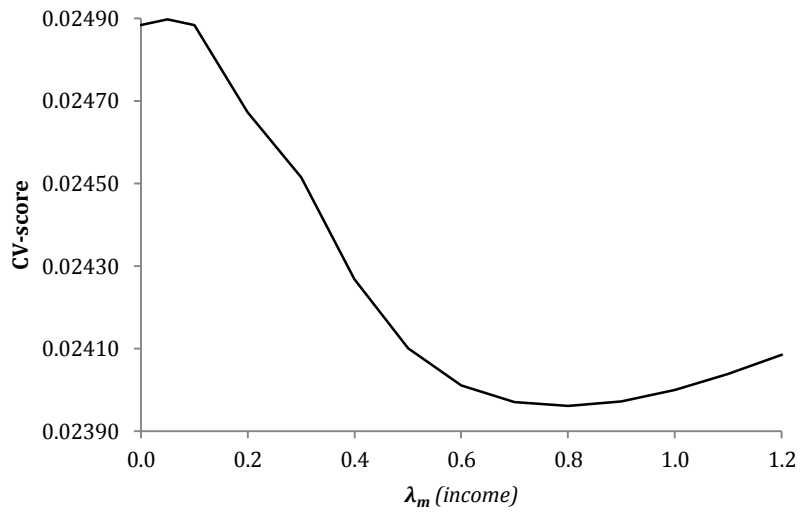


FIGURE E3 — BANDWIDTH SELECTION FOR SPECIFICATION (12)

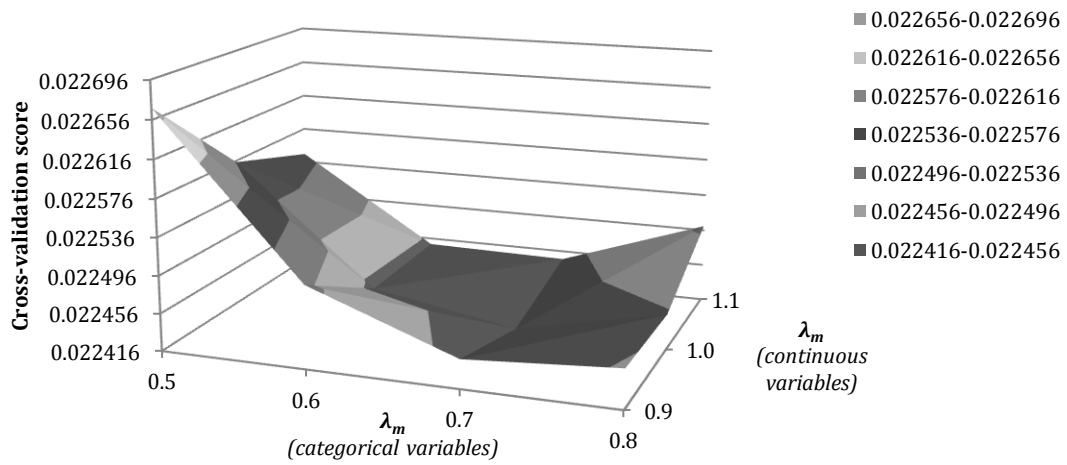
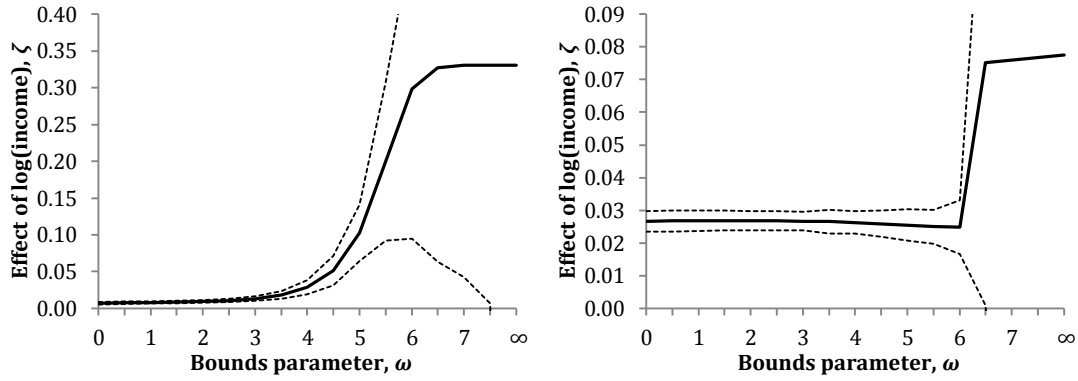


FIGURE E4 — BANDWIDTH SELECTION FOR SPECIFICATION (13)

Appendix F. Specification (13-2) for different ω



(A) WTP FOR CONSERVATION AREAS α_i^*

(B) WTP FOR LISTED BUILDING β_i^*

FIGURE F1 — EFFECT OF LOG(INCOME) ON α_i^* AND β_i^* IN SPECIFICATION (13-2)

Spatial Economics Research Centre (SERC)

London School of Economics
Houghton Street
London WC2A 2AE

Tel: 020 7852 3565

Fax: 020 7955 6848

Web: www.spatial-economics.ac.uk

SERC is an independent research centre funded by the Economic and Social Research Council (ESRC), Department for Business Innovation and Skills (BIS) and the Welsh Assembly Government.