Landscape and Urban Planning 207 (2021) 104016

Contents lists available at ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

Research Paper

Who demands peri-urban nature? A second stage hedonic house price estimation of household's preference for peri-urban nature

Cathrine Ulla Jensen [b], Toke Emil Panduro [a,*], Thomas Hedemark Lundhede [b,c,e], Kathrine von Graevenitz [d], Bo Jellesmark Thorsen [b,c]

[a] Department of Environmental Science, Faculty of Technical Sciences, Aarhus University, Frederiksborgvej 399, 4000 Roskilde, Denmark
[b] Department of Food and Resource Economics, Faculty of Science, University of Copenhagen, Rolighedsvej 23, 1958 Frederiksberg Copenhagen, Denmark
[c] Centre for Macroecology, Evolution and Climate, University of Copenhagen, Rolighedsvej 23, 1958 Frederiksberg Copenhagen, Denmark
[d] Department of Environmental and Resource Economics, Centre for European Economic Research (ZEW), P.O. Box 103443, 68034 Mannheim, Germany
[e] Centre for Environmental Economics and Policy in Africa, Department of Agricultural Economics, Extension and Rural Development, University of Pretoria, Pretoria, South Africa

HIGHLIGHTS

- Households' preferences for urban nature revealed by their choice of housing.
- A second stage hedonic housing price analysis using a single market.
- The relation between demographics and willingness to pay (WTP) is not constant across the WTP-distribution.
- Willingness to Pay for peri-urban nature increases with income and wealth.

ARTICLE INFO

Keywords:
Public policy
Progressive distribution
Green space
2nd stage hedonic regression
Quantile regression

ABSTRACT

We analyze housing markets in a suburb north of the Danish capital Copenhagen. There the concentration of affluent households decreases rapidly with distance to nature. This indicates systematic differences in preferences that are highly correlated with demographics. We assess if and to what extent this is the case by conducting a second-stage hedonic house price study, where we recover household-specific preferences for the availability of peri-urban nature. Preference parameters are identified locally through restrictions on household utility functions. We assess the relationship between demographic factors and household Willingness To Pay for nature availability across the Willingness To Pay distribution. The results of the analysis show that households paid on average 968 EUR per year for the level of peri-urban nature they have available from their home. In extreme cases, some households paid more than seven times the average amount.

Willingness To Pay for peri-urban nature increases with income and wealth across the entire population. However, an increase in income will increase Willingness To Pay more for households at the lower end of the willingness to pay distribution, compared to the effect at the high end of the Willingness To Pay distribution. Furthermore, increases in education levels are related to higher willingness to pay in the middle of the willingness to pay distribution, whereas further education has less impact on households with high willingness to pay. Single parent status, car ownership, and other factors also affect Willingness To Pay for nature availability.

Our study contributes to the discussion of the distributional aspects of environmental policies with results based on a revealed-preference method where households face a real and binding budget constraint. Our results show that the socio-economic distribution of changes in amenity values is a relevant factor to consider when evaluating policies that affect the provision of nature close to urban areas.

* Corresponding author.
E-mail address: toke@envs.au.dk (T.E. Panduro).

https://doi.org/10.1016/j.landurbplan.2020.104016
Received 28 April 2020; Received in revised form 24 November 2020; Accepted 25 November 2020
Available online 6 December 2020
0169-2046/

1. Introduction

The role of urban green spaces such as peri-urban nature areas and urban parks for urban households has received considerable attention in urban planning debates and research. These areas provide a range of services, including recreational opportunities and amenity values; they also, potentially, have various effects on health, community building, and other social values. These values are mainly, but not exclusively, enjoyed by households living near urban parks and or nature areas in the vicinity of urban areas.

Research has taken several different approaches to study the values of green space for humans. Studies have investigated the role of seeing and visiting nature areas on aspects of human health (Willis and Crabtree, 2011; van den Berg, Koole, & van der Wulp, 2003), and others have developed tools to assess the broader sets of social values associated with urban nature areas (Tyrväinen, Mäkinen, and Schipperijn, 2007). Hedonic house price studies have attempted to capture the economic use value of proximity to various forms of peri-urban nature areas using hedonic house price approaches (Sander, Polasky, & Haight, 2010; Tyrväinen and Miettinen, 2000; Lake, Lovett, Bateman, & Day, 2000; Panduro and Veie, 2013). Studies have found that the availability of green spaces has significant positive benefits for surrounding households, but have provided less insight into how these benefits are distributed in the population.

Our study contributes to the literature by modeling the heterogeneity of household preferences for nature areas. The nature areas in focus are larger continuous spaces containing natural vegetation, like open grassland, tree cover, and lakes, and with small gravel roads and walking paths, but otherwise few recreational facilities. Specifically, we go beyond the hedonic first stage implicit prices and estimate preference parameters following Bajari and Benkard (2005) in assuming that households have a declining marginal utility of goods. We thus impose a functional form and estimate households' Willingness To Pay (WTP) function based on the hedonic price function (Rosen, 1974). Through the use of quantile regression, we assess drivers of heterogeneity in WTP across the WTP distribution. We use the estimated preference parameter distribution to analyze preference heterogeneity by regressing logged WTP for a one-unit increase in supply to peri-urban urban nature against observed demographics. The relationship between demographics and WTP may not be the same across the WTP distribution, and to allow for such heterogeneity, we apply a quantile regression model.

The main obstacle in a classic second stage hedonic analysis is the endogeneity of the implicit prices obtained from the first stage hedonic price model. The literature has mainly relied on the first stage implicit prices and limited welfare analysis to marginal and localized changes as suggested by (Bartik, 1988). Using first stage implicit prices requires that the change evaluated does not affect the existing market equilibrium, and households' Marginal Rate of Substitution (MRS) remains constant over the change considered. In this paper, we adopt an approach originating from Bajari and Khan (2005) that relaxes the assumption of a constant MRS, but without incurring the endogeneity problems in the classic Rosen second stage. We identify preferences through functional restrictions on the utility function. Within the existing market equilibrium, this restriction identifies the WTP for changes that are non-marginal to the individual. The approach originated from Bajari and Benkard (2005) and was applied that same year

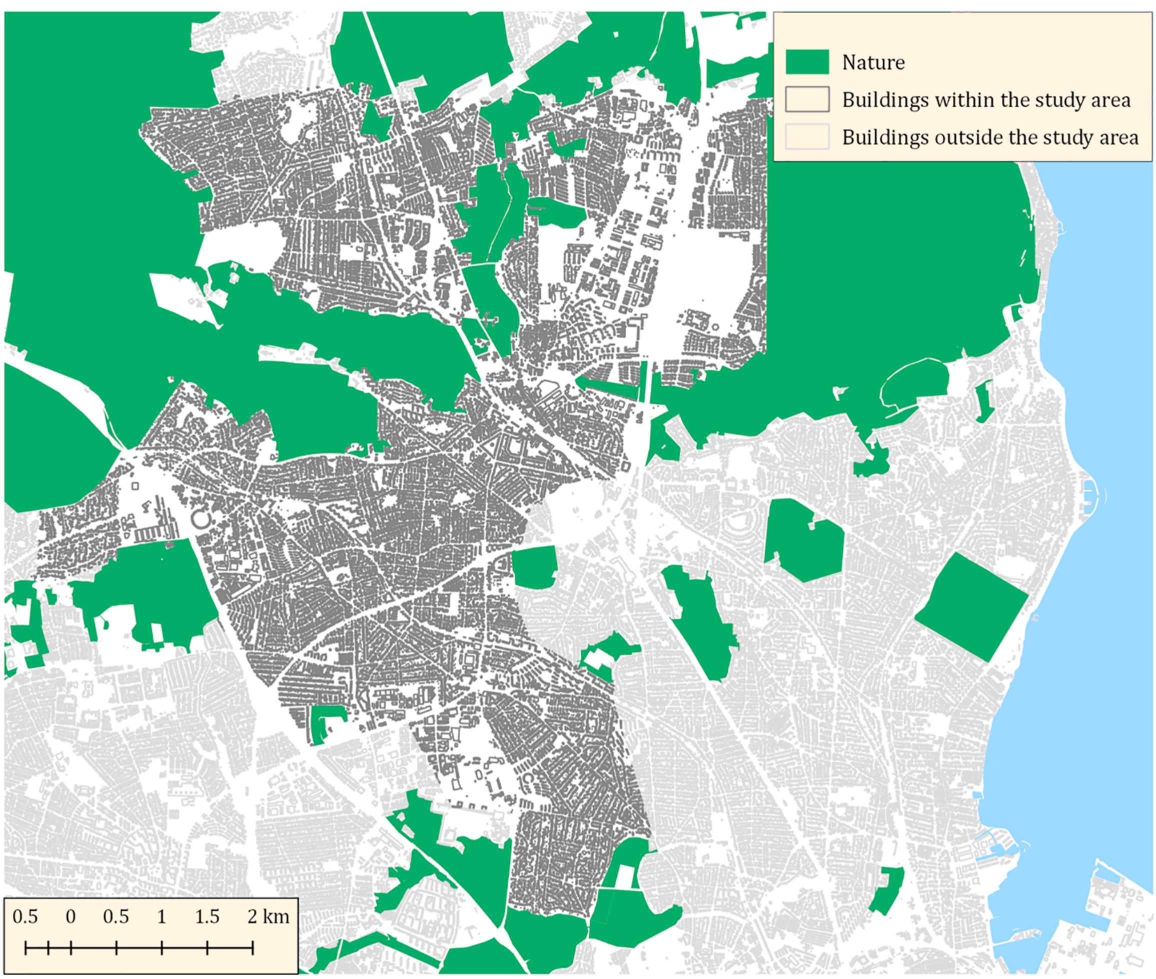

Fig. 1. The Study Area.

by Bajari and Kahn (2005), who assessed preferences for racial compositions of neighborhoods. More recent applications include von Graevenitz (2018), Panduro, Jensen, Lundhede, von Graevenitz, and Thorsen (2018), and Diamond and McQuade (2019). Chattopadhyay (1999) uses a related approach and identifies preferences through functional form assumptions in the context of air pollution.

In the environmental literature applying second stage hedonics, distributional aspects have been given little attention and treated only superficially. Brasington and Hite (2005) estimated a second-stage model for the distance to pollution hazard sites. In their first-stage hedonic regression, they included neighborhood-level measures of income, poverty, and education in the house price regression as explanatory variables. The same variables also appeared in their second stage demand regressions. They found a small but significant positive income elasticity of demand, but also found positive effects of education and number of children on the demand for distance to hazard sites. Another study focusing on urban nature areas is that of Poudyal, Hodges, and Merrett (2009), who analyzed the demand for urban green space in Roanoke, Virginia. They analyzed distributional aspects of urban green space preferences and found a significant but fairly small positive income elasticity of demand; they also found that other socio-demographic demand shifters did matter. Netusil, Chattopadhyay, and Kovacs (2010) estimated second-stage models of the demand for tree canopy cover and found weak effects of demographic demand shifters, including income. Using a related approach, namely a horizontal sorting model, Klaiber and Phaneuf (2010) examined preference heterogeneity for open space. They find evidence of substantial heterogeneity both across types of green space and types of households. In their study, wealthier households were more likely to locate near natural areas, and less wealthy households were more likely to locate near agricultural land. Tuffery (2019), also addressed the issue heterogeneity and estimated WTP for forest recreation using a random bid model and found WTP differences between affluent and older households relative to less affluent and younger households.

In general, related hedonic studies have only evaluated effects at the mean of the distribution. However, several stated preference studies have applied quantile regressions to analyze distributional aspects of environmental policies. Belluzzo (2004) used contingent valuation to estimate the WTP for the management and improvement of an important Brazilian river basin near Rio de Janeiro. He included age, income, and education as explanatory variables and noted significant differences between the size and significance levels of coefficients at the tails of the distribution, suggesting that the respondents who would benefit from the project differ significantly in terms of socio-demographic characteristics from those who would experience a welfare-loss. In their study of WTP for air and noise pollution reductions via the introduction of hydrogen buses in London, O'Garra and Mourato (2007) also find that determinants of WTP vary across the WTP-distribution. Quantile regressions have also been used to analyze the distributional aspects of changes in urban nature by Notaro and De Salvo (2010), who assessed tourists' WTP – using contingent valuation – to save an urban forest near Garda Lake in Italy, where a tree species is threatened by diseases. They found that the lower median of WTP was unaffected by income, whereas the level of education was only a factor at the low end of the distribution.

Our study contributes to the literature on the value of peri-urban nature areas by exploring preference heterogeneity and WTP variation using quantile regression. The hedonic approach has the strength of being based on actual market behavior made by households with a binding budget constraint, as opposed to the stated preference literature. Results show that WTP for nature availability increases with income across the entire WTP distribution in the population. Furthermore, higher education is related to higher WTP at the median, whereas education has less impact on households with extreme values of WTP. The presence of household members above the age of 61 years is associated with an increase in WTP for those households with an otherwise low WTP. Overall, household heterogeneity contributes not only to explaining levels of WTP but also to explaining dispersion in the WTP distribution.

Table 1
Variable description.

	Description
Area	The size of the living area in square meters.
Toilets	Number of toilets
Garden	The size of the garden in square meters
Roof: tile	Dummy variable that describes whether the roof is made of tile. 1 corresponds to being built of tile and 0 corresponds to not made of tile.
Roof: Cement	Dummy variable that describes whether the roof is made of cement. 1 corresponds to being built of cement and 0 corresponds to not made of cement.
Bathrooms	Number of bathrooms
Rebuilt in 70-ies	Dummy variable that describes whether the dwelling has undergone major renovation during the 1970 s. 1 corresponds to major renovations and 0 corresponds to no major.
Rebuilt in 00-s	Dummy variable that describes whether the apartment has undergone major renovation during the 2000 s before the apartment is sold. 1 corresponds to major renovations and 0 corresponds to no major renovation.
x	The × coordinate measured in UTM 32 North WGS 1984
y	The y coordinate measured in UTM 32 North WGS 1984

Table 2
Summary statistics of the variables in the hedonic model.

Statistic	Unit	Mean	St. Dev.	Min	Max
Price	1000 EUR	452.474	140.150	103.306	897.541
Area (log)	m^2	4.772	0.265	4.094	5.704
Toilets	#	1.560	0.581	1	4
Garden (m^2)	m^2	646.192	358.116	0	2763
Large Road	m	111.738	131.960	0	386.803
Nature density	Hectares within 800 *m*	29.477	31.228	0	137.907
Tile roof	1/0	0.513	0.500	0	1
Cement roof	1/0	0.047	0.211	0	1
Bathrooms	#	1.251	0.479	0	3
Rebuilt in 70-ies	1/0	0.096	0.295	0	1
Rebuilt in 00-s	1/0	0.043	0.203	0	1

2. Material and methods

2.1. Study area and housing transactions

The study area covers 3900 ha of land just north of the Danish capital Copenhagen. The area is characterized by large nature areas, with old forests, lakes and open meadows interspersed with urban areas (see Fig. 1).

The dataset consists of 2376 single-family detached properties traded at arms-length between 2007 and 2010. A total of 72 properties, traded for more than 900,000 EUR or <100,000 EUR, were removed as they seemed to represent potential measurement or reporting errors (ca. 3% of the sample). For example, the lower bound most likely excludes errors or cases where the house is in a poor condition, and the property is traded as a building plot with a negative scrap value of the buildings. The upper bound excludes errors and very luxurious homes, which may have special unobservable characteristics. The impact on the first stage results from excluding these homes is marginal. Full estimation results, including these 72 properties, are found in Table A3 in the appendix. In addition to the sales price, the date of the transaction, and the exact geocode of the property, the dataset also includes the structural characteristics of the property, e.g., size of the living area, etc. This information was extracted from the Danish Registry of Buildings and Housing (SKAT, 2012). Spatial variables, which capture various qualities of the property's surroundings, were calculated for each property using R (R

Core Team, 2015) and ArcGIS 10.2.1 (ESRI, 2011 2015). We control for border effects of the study area delineation by allowing spatial variables to include areas outside the delineation according to their definition. The spatial data were supplied by the Danish Geodata Agency (The Danish Geodata, 2011) and by the Danish Business Authority (Danish Business Authority, 2011). Please see Tables 1 and 2 for summary statistics and variable descriptions.

2.2. Defining nature areas and their availability

In this paper, we distinguish between nature areas and other types of urban areas. In general, green space is not a uniform good, but rather several distinct goods that enable different uses (Panduro and Veie, 2013). Within the survey area, green space was classified into different categories. Peri-urban nature areas are the focus of our study. These were identified as being large continuous spaces containing open grassland, tree cover, and lakes, and containing small gravel roads and walking paths, but otherwise few recreational facilities. These features of the nature area enable people to move through the landscape along the gravel roads or walking paths. People perceive nature areas to be more natural landscapes that are not maintained by society (Vining, Merrick, and Price, 2008). This feature distinguishes peri-urban nature areas from other types of green spaces such as urban parks, which are more carefully groomed, common areas between buildings, churchyards, and sports fields as defined by Panduro and Veie (2013). In our suburban study area, there are very few green spaces falling into the park category, and hence very few traded homes within proximity of a park. In contrast, the supply of "peri-urban nature" areas is high in this market and makes it well-suited for a study of WTP for this type of green space.

In the hedonic literature, proximity and density measures have often been used to capture access to nature areas as a part of the housing bundle, see, e.g., Tyrväinen and Miettinen (2000), Kong, Yin, and Nakagoshi (2007), or Orford (2002). Different variants of density measures have been applied in previous studies. Density and patchiness measures of urban green space were used by Kong et al. (2007) in combination with distance measures. Studying the value of peri-urban forest land areas in North Carolina, USA, Cho, Jung, and Kim (2009) applied patch size and patch and edge density measures for both deciduous and evergreen forests. Mansfield, Pattanayak, McDow, McDonald, and Halpin (2005) and Jiao and Liu (2010) also applied forms of density measures, whereas absolute size measures of the nearest green spaces are applied by Morancho (2003).

We measure nature availability in the housing bundle using a density measure in hectares. For each dwelling, we calculate how much area is taken up by peri-urban nature areas within 800 m of each house. The 800-meter truncation, measured as the Euclidian distance, corresponds to<15 min walking at a speed of 5 km/h which. The distance truncation corresponds to the findings in forest recreation literature on typical (median) time people spend getting to a recreational area (Jensen, 2003). We tested several distance bands as well as network distances and found similar results. The choice of Euclidean distance over network distance was mainly due to potential errors in the road network that potentially would produce more bias than accuracy. In the dataset, seemingly connected roads are in fact, not connected, resulting in unrealistic travel distances.

We estimated the first stage hedonic price function with more than 25 control variables and using school districts as a fixed effect. A short description and descriptive statistics is presented in Tables 1 and 2.

2.3. Socio-demographic data

We used socio-demographic variables to decompose household-specific preferences for access to nature areas. Each property was linked to socio-demographic data from Statistics Denmark. The socio-demographic data describe the household occupying the property in 2011 using a number of relevant variables, such as income, education, car-ownership, etc., cf. Table 3. Due to the sensitive nature of individual-level socio-demographic data, they were spatially blurred using a raster mosaic of 100*100 m, which was subsequently refined and matched to individual properties by Geomatic A/S. The descriptive statistics of the socio-economic variables are presented in Table 4 as the mean value of distance-subgroups, divided into distance bands of 200 m. The table shows, among other things, that the mean income, wealth and share of households with more than one adult decreases with distance to the nearest nature area.

Table 3
variable description of the socio-economic variables in willingness to pay function.

	Description
Income (1000 EUR)	Measure the income of the highest-earning person in the household based on the household's tax-reports
Wealth (1000 EUR)	Measure of wealth, that includes all assets and passives except if the household owns a share of a home in a housing cooperative. Based on the household's tax-reports
Higher education	The adults in the household have a minimum of 5 years of higher education. The variable is constructed as a dummy variable where 1 equals have a minimum of 5 years of higher education and 0 equals having<5 years of education.
Self-employed	The variable is constructed as a dummy variable where 1 equals having a self-employed person in the household and 0 equals not having a self-employed in the household
Outside the workforce	The variable is constructed as a dummy variable where 1 equals having an adult outside the workforce in the household and 0 equals not having an adult outside the workforce in the household
Top manager	The occupation of a member of the household is top-manager. The variable is constructed as a dummy variable where 1 equals having a top manager in the household and 0 equals not having a top manager in the household.
Car owner	The household owns a car. The variable is constructed as a dummy variable where 1 equals owning a car and 0 equals not owning a car.
Single	The household contains no more than one adult. The variable is constructed as a dummy variable where 1 equals having no more than one adult in the household and 0 equals having more than one adult in the household.
Single parent	The household consists of only one adult and one or more children. The variable is constructed as a dummy variable where 1 equals having one or more children in the household and 0 equals having no children in the household.
Above age 60	The oldest member of the household is above 60 years of age. The variable is constructed as a dummy variable where 1 equals having at least one adult above 60 years of age in the household and 0 equals having zero adults over 60 years of age in the household.

2.4. Hedonic pricing and welfare changes

The hedonic method was first described in the seminal paper by Rosen (1974). The first stage has been a workhorse in valuation studies ever since. In short, the first stage of the hedonic pricing method identifies marginal implicit prices by regressing housing prices on housing attributes. Housing attributes are all those characteristics that are valued by buyers and sellers and thus affect the market value of the home. This includes attributes of the property itself, e.g., number of rooms or build-year and attributes of the surroundings, e.g., school quality, access to green space and nature areas.

Formally the hedonic price function h maps the relationship between the price p_i of property i, our focus variable of peri-urban nature areas, q_i, and K other housing characteristics captured by the vector X_i:

$$p_i = h(X_i, q_i) \tag{1}$$

Rosen (1974) showed that just from the hedonic pricing function, we can obtain estimates of the WTP for marginal changes in attributes, often referred to as implicit or marginal prices. He also described a "second

Table 4
Demographics (MEAN) given shortest distance to a nature area.

	0–200 m	200–400 m	400–600 m	600–800 m	>800 m	Full sample
Income (1000 EUR)	101.36	89.19	81.21	78.16	70.28	84.69
Wealth (1000 EUR)	507.88	402.12	349.31	333.72	280.31	379.49
Higher education	0.90	0.82	0.70	0.72	0.55	0.74
Self employed	0.17	0.15	0.26	0.22	0.27	0.21
Top manager	0.75	0.72	0.53	0.56	0.39	0.60
Employee	0.08	0.11	0.19	0.21	0.31	0.18
Age min 61	0.39	0.25	0.19	0.14	0.12	0.23
Single	0.06	0.11	0.15	0.12	0.11	0.11
Single parent	0.05	0.09	0.13	0.10	0.11	0.09
Nature availability within 800 m (ha)	68.07	46.49	18.30	3.26	0	29.48
Properties	541	506	467	350	512	2376

stage" analysis, where household demand schedules could be obtained by regressing marginal prices from the first stage on quantities. In 1988, Bartik pointed out that Rosen's second stage suffered from an inherent problem of endogeneity. The pricing of housing goods is seldom linear, and therefore a household chooses its marginal price and consumption level of a housing good simultaneously. Bartik proposed a multiple market strategy, where exogenous variation in marginal prices arising from multiple markets provides a solution to the endogeneity problem. This strategy would ideally reveal the entire demand schedule for each household type by pooling and observing household types across multiple markets. Bartik's suggestion is the most implemented second stage identification strategy in the literature. However, there are few second stage studies, perhaps because the multiple market strategy is data-intensive and empirically challenging due to other endogeneity problems. The foremost of these problems is that people may sort across markets, which means that variation across markets is no longer exogenous to preferences. In many cases, the literature has been dominated by first stage hedonic studies and marginal prices. These are in limited cases sufficient to identify an upper bound on the WTP for an improvement in the provision of an amenity (Bartik, 1988; Palmquist, 1992).

Bartik (1988) noted that a sufficiently small and localized change would not change the existing market equilibrium, and thus the households would face the same marginal prices before and after the change. Therefore, the researcher can obtain an upper bound on the total WTP for change, with just first stage implicit prices and a credible explanation of why the valued change is marginal to the market. With non-zero moving costs or other rigidities, the welfare improvement would typically be smaller under the assumption of declining marginal utility as the household's choice was optimal under the original equilibrium before the change.

To explicitly discuss the underlying assumption on the MRS, we need to specify household utility. Each household owns one house, and so we use the subscript i for both households and homes. A household i receives utility U_i by occupying house i and consumes the housing goods X_i and q_i and all other goods captured in the numeraire Hicksian composite good c_i.

$$U(X_i, q_i, c_i) \tag{2}$$

In a market in equilibrium, all households occupy a home and are unable to change their consumption of a housing attribute without reducing their overall utility. Therefore, if a household i buys the home i with the peri-urban nature q_i, its MRS between any alternative consumption c_i and q_i must be equal to the marginal cost of q_i:

$$\frac{\delta U(X_i, q_i, c_i)/\delta q_i}{\delta U(X_i, q_i, c_i)/\delta c_i} = \frac{\delta h(X_i, q_i)}{\delta q_i} \tag{3}$$

Eq. (3) states that for a housing bundle to be optimal, the MRS equals the marginal cost for each good consumed. In other words, $\frac{\delta h(X_i, q_i)}{\delta q_i}$ is the marginal implicit price for q obtained from the first stage hedonic regression.

With the declining marginal utility, the left side of the above expression will decrease as q_i increases, which implies that households have a positive WTP for an increase in q_i, but decreasing with the amount of q_i already available. The utility function adopted by Bajari and Benkard (2005) satisfies this property by modeling utility as an additively separable concave function that is linear in income. These assumptions result in a MRS that is specific for each housing good:

$$U_i = \sum_k \gamma_{ik} \log(x_{ik}) + \gamma_{iq} \log(q_i) + c_i \tag{4}$$

γ_{iq} is the household's preference or taste for q, where a higher number indicates a stronger taste for good q. The quasilinear utility function with weak separability is not common in the hedonic literature, but functions with similar properties have been used within the stated preference literature (Li and Mattsson, 1995).

Furthermore, it results in a very simple first-order condition:

$$\frac{\gamma_{iq}}{q_i} = \frac{\delta h(X_i, q_i)}{\delta q_i} \tag{5}$$

We are interested in the preference parameter γ_{iq} and rearrange (5) into (6):

$$\gamma_{iq} = q_i \frac{\delta h(X_i, q_i)}{\delta q_i} \tag{6}$$

By assuming a simple restriction on the utility function, the household's preference for peri-urban nature is estimated based on q_i, the observed level of q chosen by household i, and using the recovered preference parameter, we can calculate WTP for a change in q from q^0 to q^1:

$$WTP_{iq} = \gamma_{iq} \log\left(\frac{q^1}{q^0}\right) \tag{7}$$

It is useful to think of the identifying assumption on utility as a local approximation for a given level of income and market equilibrium. As a result, it is also most likely to provide accurate insights into changes that do not depart too far from the observed choices made. The bigger the change under analysis, the larger is the role played by the functional form assumption in determining WTP for that change. Because this method introduces declining marginal utility of the good, it will always result in a smaller WTP than one based simply on linearly rescaling the first stage implicit prices.

3. Empirical application

3.1. First stage pricing function

There are two primary concerns when estimating the hedonic price curve: the functional form and the risk of omitted variable bias. Here we examined the functional form for each characteristic by estimating different models and examining the fit. To address the second concern, we report three different models using three different approaches to control for unobserved trends in time and space. The price function was estimated using a general linear model (GLM) and its semi-parametric extension, a general additive model (GAM). A GLM is a general form of the classic OLS but allows the distribution to differ from a normal distribution and utilizes a link function between the dependent and independent variables. House prices are strictly positive, and therefore we use a gamma distribution and implement the semi-log through a "log-link" function.

The GLM is written as:

$$\ln(p_i) = \sum_k \beta_k x_{ik} + \beta_q q_i + t_y + \delta_l + \xi_{il} \tag{8}$$

and the semi-parametric GAM:

$$\ln(p_i) = \sum_k \beta_k x_{ik} + \beta_q q_i + f_1(t_i; S_1) + f_2(x_i, y_i; S_2) + \delta_l + e_i \tag{9}$$

In both models, q_i is the density of peri-urban nature, x_{ik} is housing characteristic k for dwelling i, β_q and β_k are the estimated parameters. In the GLM, ξ_{il} is an error-term clustered at the neighborhood level as defined by the spatial fixed effects (δ_l). In the GAM, the error term is assumed to be i.i.d. as the smooth spatial component captures additional spatial variation and clustering the errors is not feasible in the estimation.

The two models differ in their control for spatial autocorrelation and time trends in the variables included. The GLM includes spatial fixed effects (δ_l) and different transformations (t_y) of the transaction date. The GAM, in addition to the same spatial fixed effects, includes a non-parametric smooth component in time t and over the spatial coordinates (x, y) estimated by fitting one-dimensional and two-dimensional splines to the data: $f_1(t_i; S_1)$ and $f_2(x_i, y_i; S_2)$. The parameter S in the two smooth functions represents the number of spline basis functions used to fit the smooth component. The component consists of a weighted sum of individual spline basis functions (polynomials of varying degrees) to flexibly fit a smooth function to the data. The individual weights (coefficients on each spline basis function) are estimated from the data and the flexibility increases with the number of splines along with the risk of overfitting (Wood, 2017).

When deciding on how to specify fixed effects for the GLM, we calculated and tested a number of neighborhood definitions using different spatial ranges in the pricing function, as suggested by von Graevenitz and Panduro (2015). Due to the importance of school quality for many households in choosing a home, our preferred specification uses school attendance zones as the spatial scale for our fixed effect. However, the estimates were robust across a range of neighborhood definitions.

The GLM requires an assumption on what time trend specification to include in τ_y and the spatial scale of neighborhood dummies, δ_l. The spatial–temporal control is more data-driven when it is modeled non-parametrically in the GAM. In the GAM omitted spatial characteristics were controlled for non-parametrically by smoothing over space $f_2(x_i, y_i; S_2)$ using the spatial x_i and y_i coordinates of each sold property.

In our application S is recovered by using generalized cross-validation and is thus data-driven. This approach avoids making assumptions about the structure and extent of the unobservable spatial processes in the data (von Graevenitz and Panduro, 2015). The researcher does choose the maximum degree of flexibility by choosing the number of spline basis functions. Sensitivity analysis of this assumption can easily be made and corresponds to analyzing the impact of using different spatial fixed effects or to assume different neighborhood matrixes in a parametric spatial econometric model. The advantages of this semi-parametric approach are starting to be recognized in the hedonic literature, and a few recent applications are found in Cajias, Fuerst and Bienert (2016) and Fritsch, Haupt, and Ng (2016).

3.2. From implicit prices to household-specific preferences

The following describes in detail how the preference parameter is calculated based on Eq. (6). We model the choice of housing as a static problem. This assumption is quite strong, though it is standard in the hedonic literature. Bishop and Murphy (2011) show that the assumption of myopic consumers can lead to biased estimates, especially for attributes that are expected to change over time. Given the Danish urban planning regulation and the urban development history of the study area, the supply of nature areas has been stable over a long period, and we believe any resulting bias to be limited.

The observed trade prices, p_i, were first converted to 2011 prices ($\widehat{\mathrm{P}}_{\mathrm{i},2011}$), using the price trend in the data, and then to a perpetual annuity by assuming perpetual life for the house asset and multiplying the observed price with a discount rate π, using a rate of 3%. This approach is also used in Day, Bateman, and Lake (2007) to convert house prices to an annual rent and shown by Phaneuf and Requate (2016). The house prices are therefore expressed as a perpetual annuity or an "annual rent" in 2011 prices. To get the annual marginal WTP in EUR, we multiplied this with the implicit price from the first stage, $\widehat{\beta_{\mathrm{q}}}$:

$$\widehat{\beta_{qi,\,annual}} = \widehat{P}_{\mathrm{i},2011} \times \pi \times \widehat{\beta_q} \tag{10}$$

Then we multiplied this annual payment with the chosen level of q in the housing bundle that household i chose, q_i based on the first order condition in Eq. (6):

$$\widehat{\gamma}_{\mathrm{iq}} = q_i \times \widehat{\beta_{qi,\,annual}} \tag{11}$$

The preference parameter γ_{iq} is therefore expressed in 2011 EUR per year.

As the identification of preferences is based on the equality between the MRS and the implicit marginal price (Eq. (3)) we cannot use this method to identify preferences for households who chose a corner solution, i.e., no access to peri-urban nature areas. In total 512 households (around 21% of the sample) bought a home with no peri-urban nature within 800 m. For these households, the MRS for the good is smaller or equal to the marginal cost. Without further assumptions about the distribution of the taste parameter, we cannot recover γ_{iq} for these 512 households, but only conclude it is bounded between 0 and the corresponding γ_{iq} calculated for these households from their first-order condition. We, therefore, exclude these 512 households from the further analysis of preference heterogeneity.

3.3. Analysing the variation in WTP

We analyzed preference heterogeneity by regressing estimates of WTP on demographic variables, allowing for different effects across the WTP-distribution. A conditional mean model implicitly assumes that the effect of covariates moves the entire distribution by a fixed factor. One of the main contributions of this paper is to relax that assumption by, in addition to standard OLS, also estimating quantile regressions (QR) to examine differential impacts across the conditional WTP distribution.

We estimated:

$$\mathrm{Q}_\tau(\ln(\mathrm{WTP})|\mathrm{D_i}) = \alpha_\tau + \sum_{\mathrm{d}}^{\mathrm{D}} \alpha_{\mathrm{d}\tau} \mathrm{D_{di}} \tag{12}$$

where D_i are socio-economic covariates, α_τ is an intercept that captures the intercept for the τ^{th} decile, D_{di} are D observed demographic characteristics, and $\alpha_{d\tau}$ is a vector of parameters describing the variation that can be explained by each observable characteristic in the τ^{th} decile. The coefficients in α_d represent the marginal effect of the explanatory variable d on the τ^{th} conditional quantile in the estimated preference distribution. We estimated QRs in R using the *Quantreg* package (Koenker, 2013). Note that when fitting a regression conditional on τ, all observations contribute to the fitting of the regression but with different

Table 5
Estimated implicit prices for nature availability.

Spatial control	Fixed effect	Fixed effect and spatial smoothing	Fixed effect and clustered residual
Time control	Time smoothing	Time smoothing	Polynomial dates
Model	GAM	GAM	GLM
	(1)	(2)	(3)
Nature availability	0.00182***	0.00172***	0.00187***
(ha within 800 m)	(0.00018)	(0.00019)	(0.00023)
Constant	3.86384***	3.76895***	3.86384***
	(0.09273)	(0.09196)	(0.14599)
Adjusted R^2	0.544	0.582	0.544
Log Likelihood	−14,124	−14,102	−14,124
AIC	28,222	28,204	28,231

Note: N = 2376. Std. errors in parantheses. Significance levels are: *p < 0.1; **p < 0.05; ***p < 0.01.

Table 6
Preference-parameter.

	Min	1st quartile	Median	Mean	3rd quartile	Max
Preference parameter	0.0001	189	674	968	1491	7128

Note: N = 1864. Values are EUR/year.

weights depending on their location above or below the quantile in question. This makes quantile regression less sensitive to extreme observations and is useful when it comes to studying asymmetric distributions. (Koenker and Hallock, 2001).

4. Results

4.1. Recovering tastes and estimating WTP

We present three different versions of the hedonic price schedule with varying temporal and spatial controls. The results are shown in Table 5, where we only list estimates of the parameter for the peri-urban nature area availability variable, the intercept, and model statistics. The model includes more than 25 explanatory variables, and full results are found in Appendix Tables A1 and A2. The coefficients of the control variables in the models conform to expectations, e.g., increasing the size of the living area or the number of bathrooms is associated with a higher price, whereas decreasing the distance to a larger road is associated with a lower price.

The parameter for the availability of peri-urban nature areas reflects a marginal implicit price of an increase in density of nature areas corresposnign to one additional ha within an 800 m radius of the property. It corresponds to a price premium of between 0.17% and 0.19% per house per ha, which again corresponds to a marginal WTP of approximately 800 EUR per additional ha and household for an average-priced property in the sample (the mean in the sample is 450,000 EUR (2011)).

We control for omitted spatial covariates by imposing a spatial fixed effect on school districts (model 1), by including both a fixed effect and a smoothing term across space (model 2) and by using a spatial fixed effect where the residuals are clustered (model 3), thereby taking into account that observations within a school district are more correlated than those between school districts. We find our parameter estimates related to nature availability to be robust to a range of spatial controls using fixed effects on different spatial scales (municipality, postal code, road code and school district) and flexibility of the spatial smoothing term. This lends support to our assumption that unobserved neighborhood quality is uncorrelated with the observed quality we control for.

Using model (2) of Table 5, we calculate the resulting distribution of household-specific preference parameters and summarize them in Table 6 in the form of annual measures, using a discount rate of 3%, cf

Table 7
Annual WTP for a change in nature density (EUR/year).

	Min	1st quartile	Median	Mean	3rd quartile	Max	St. Dev.
10–30 ha	0	207	741	1063	1638	7831	1036
20–30 ha	0	77	273	392	604	2890	382
30–60 ha	0	131	467	671	1033	4941	654
40–60 ha	0	77	273	392	604	2890	382
50–60 ha	0	34	123	176	272	1300	172
Implicit price		24 EUR/ per ha/per year/per household					

Note: N = 1864, the implicit price is based on an average priced property in 2011 at a 3% discount rate.

Eq. (10) above. We find a median preference parameter for the 1864 households and who bought nature as a part of their housing bundle of 674 EUR/per year. The distribution is right-skewed with a few households showing a very high preference for nature access.

The WTP for a change in nature density from q^0 to q^1 can be calculated as $\gamma_{iq}\log\left(\frac{q^1}{q^0}\right)$, using the assumed utility function and the recovered preference parameter. Table 7 summarizes the household-specific WTP for a selected set of discrete changes in the availability of peri-urban nature within the 800 m radius. The median density for the households who bought nature density is approximately 30 ha, the lower quartile is approximately 20 ha, and the upper quartile is approximately 60 ha.

The restriction on the functional form of the utility function implies that the WTP for a percentage change in nature density is constant, exemplified in the table with a change from 20 to 30 ha and a change from 40 to 60 ha, which both corresponds to the same relative increase.

4.2. Variation in WTP across demographics

The log of WTP is regressed on socio-demographic variables using OLS and quantile regression. The quantile regression approach offers an insight into the heterogeneity across the price distribution, but it comes at the cost of increased complexity. The quantile regressions were run on each 5-percentile from the 10th (0.1) to the 90th (0.9) percentile, resulting in 17 estimated coefficients for each parameter for each model run. Therefore, we present results graphically (see Fig. 2). The full results for the 20th, 40th, 60th, and 80th percentile are reported in Table A2 the appendix. The signs of the estimated parameters conform to expectations, and when significant, they have the same sign, but their magnitude differs across the distribution.

The baseline household behind Fig. 2 has an annual income of 81,129 EUR and a wealth of 376,846 EUR, corresponding to the medians in the sample. The household consists of a minimum of two adults, and at least one household member is an employee; no one is over age 60 and no member holds a university degree. The dependent variable is logged WTP, and the estimated parameters are thus the percentage effect on WTP for a one-unit increase in the variable of interest holding all else constant. The OLS estimates show the impact of the covariate on the conditional mean WTP. In Fig. 2, the OLS parameter estimate is presented in red using a full line, and confidence intervals are shown by two dotted lines. The quantile regression coefficient estimates for each 5-percentile are connected by a black dotted line, and the confidence intervals are outlined by the grey areas in the figure.

The OLS estimate associates an increase in income of 1000 EUR with an increase in WTP of 0.9%. Turning to the quantile regression results, the increase ranges from 0.6% to 1.2%, with the largest increase in the lowest part of the WTP distribution. However, as the confidence intervals for the OLS and the quantile regressions overlap, there is no significant difference in the relationship between income and WTP across the distribution of WTP.

Wealth exhibits the same profile as income, with the highest relative

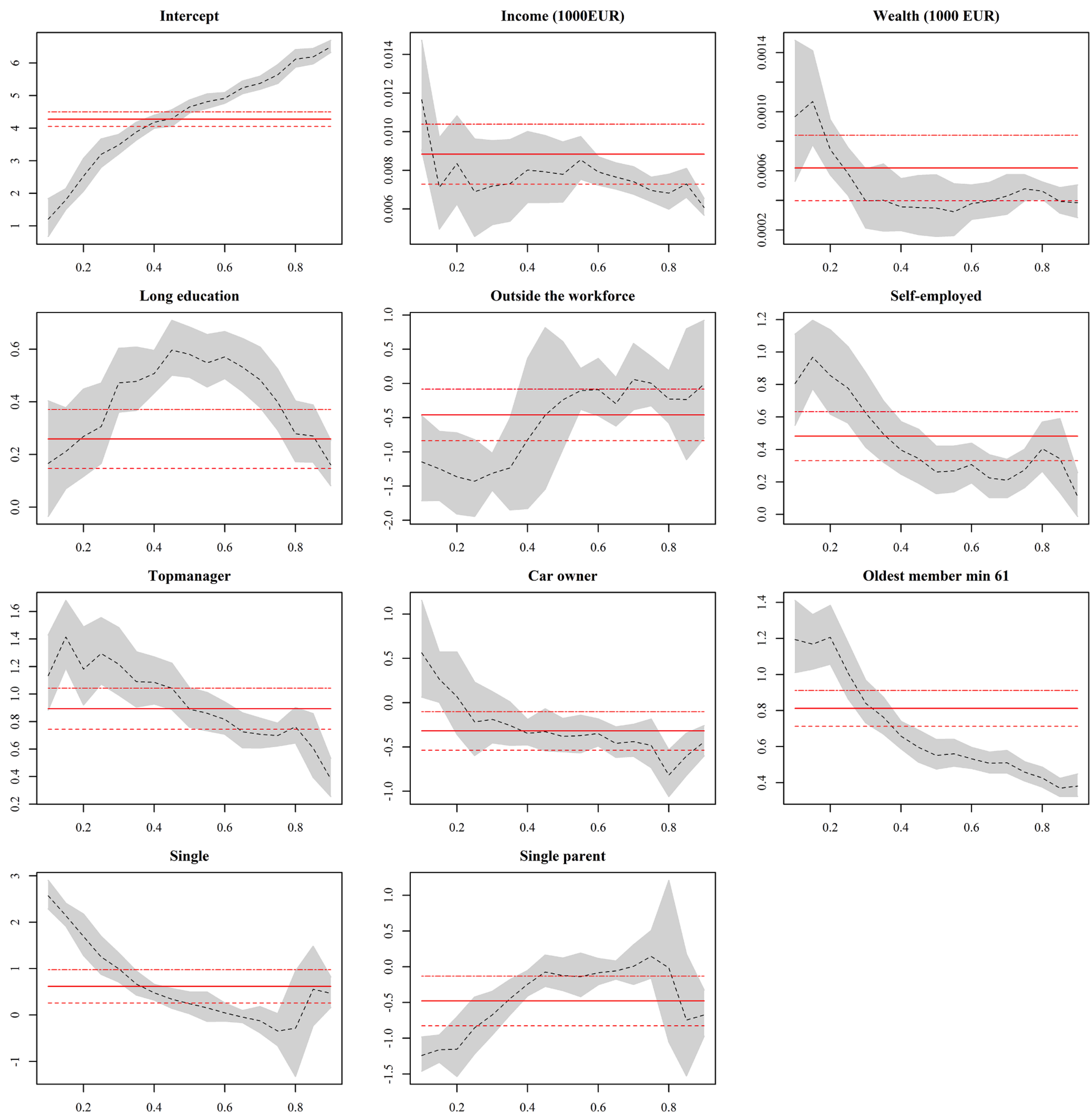

Fig. 2. Results of explaining WTP by observed demographics. Note: the y-axes describe the parameter estimate of the quantile regression, and the x-axis describes the corresponding quantiles of the WTP distribution. The fat red line illustrates the mean parameter estimate while the dotted red line represents the 95% confidence level for the mean estimates. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

effect in the low part of the distribution and a relative constant effect across the middle-to-high end of the distribution. Note that the analysis here is descriptive. Thus, we cannot say anything about what the individual household would do should their income increase; but, through the quantile regression, we can see the degree to which the relationship between income and WTP varies for households with high vs. low WTP for the availability of peri-urban nature.

On average increasing, the education for one member of the household is associated with an increase of 25% in WTP. The quantile regressions are conditional on the quantile so that the estimated parameter gives the change in WTP associated with increasing the education level, assuming that the position of the household in the WTP distribution among all other households with the same characteristics does not change. In a regression conditional on the 40th percentile, if the household has the 40th (0.4) percentile WTP and one member changes education levels, the WTP will increase by 53%, corresponding to the 40th (0.4) percentile for households where a minimum of one member has higher education. In absolute terms, this increase corresponds to 36 EUR/year.

Higher education is associated with higher WTP across the whole distribution, but matters the most around the median WTP and less so in the tails. Being outside the workforce is associated with lower WTP for households whose WTP is below the median. For households with a WTP above the median, a change from employment to a situation where all members are either unemployed, retired or students does not affect their WTP. For both households with self-employed members and top-managers we find a light tendency that WTP is higher among households with otherwise low WTP, however, again the confidence intervals overlap with those of the OLS estimate suggesting that the difference across the distribution is not significant.

Older households (defined by the oldest household member) show a higher WTP compared to households with no members above the age of 60. The impact ranges from 120 to 40%, with the highest impact on low WTP. This suggests that as the household ages, the dispersion in WTP is lower, with a higher increase in relative terms happening on the left side of the WTP-distribution.

There are very few single households and even fewer single households with children. Up to the median WTP, we observe that single households exhibit a higher WTP compared to households with two adults and a minimum of one child. For single parents, the “single” and “single parent” coefficients should be interpreted together. In our sample, 10% of the households are single parents, which are households we would expect to have a lower disposable income compared to a household consisting of two adults only, as the income measure is the yearly income of the adult with the highest income. Even so, we find that in the lower parts of the WTP-distribution single parents have a WTP which is approximately 50% larger (-1.154 + 1.693 = 0.539) than a similar household with two adults. Car ownership could affect WTP negatively because a car reduces the time cost of transport (which makes it easier to reach substitutes for nearby nature areas) and reduces disposable income. This is the case on average and particularly for households with a high WTP. However, it is not so for households with a low WTP, where instead we see that car-ownership affects WTP positively, suggesting that at this end, households tend to buy more nature when they own a car, for example, by increasing the distance to work.

5. Discussion

5.1. Who demands peri-urban nature?

We find that the implicit price for the availability of peri-urban nature areas is positive and that property prices increase with this variable. Eliciting preference parameters, we find that variation in WTP for the availability of peri-urban nature areas can, to some degree, be explained by demographics. Across the whole WTP distribution, WTP increases with wealth and income, weakly supporting the hypothesis that nature availability is a good with a progressive benefit distribution. For households with a low WTP for nature areas, the main determining factors for their WTP are employment status, whether they are reaching retirement age and whether they are single parents. In contrast, households with a high WTP do not have higher WTP if they are single parents, but their WTP differs more by car-ownership than for the rest of the population. The few single parents who have a high WTP may be different in other unobservable characteristics compared to those with a low WTP. Households with a high WTP for nature areas may also have a preference for the good, which overshadows the effect on disposable income from being a single parent.

The result of the analysis is somewhat similar to the few studies that focus on preference heterogeneity. Tuffery (2019) finds a positive WTP for forest areas among affluent and older households using a random bid model based on housing data. However, Tuffery (2019) interestingly also find that less affluent households have a negative willingness to pay for forest area. We find no evidence to support such conclusions in this paper. Other related methods using choice experiment analysis find differences in WTP for nature and urban green spaces due to age, affluence, and family structure (Ta, Tardieu, & Lev, 2020, Tu, Abildtrup, & Garcia, 2016). Preference heterogeneity may also exist in relation to different types of green space and their recreational potential (De Valck et al., 2017). At present, studies are too few, and their results diverge. There is a need for more studies that explore preference heterogeneity and that cover different types of green space and nature.

5.2. Caveats and key assumptions

We can explain 10–16% of the preference variation in our quantile regression model of WTP. In the context of microdata and given that in the estimation of the preference parameter, we include all unexplained variation; this is a fairly high explanatory power. Even so, a large percentage of the variation in preferences is left unexplained, which could also imply that classical demographic variables – or at least those used in this paper – do not fully capture preference variation. Our data on demographics describe residents in 2011. Thus, we essentially assume that the new residents moving to the area during the period 2007–2010 have the same observed demographics as those who lived there in 2011 – or at least that the mean demographic parameters of the 100 × 100 m fitted mean estimates were constant over the period. Our sample covers only four years, and, given the short period, we have no reason to believe that the composition of households in terms of demographics has changed substantially.

We have shown that the density of high-income households decreases with decreasing nature availability across our case area. This, of course, reflects that they have – on average – higher WTP for nature availability and hence sort themselves disproportionally into these areas within our housing market – as opposed to lower-income households, who, on average, have lower WTP for nature availability. Such sorting is further exacerbated by feedback effects as the quality of neighbors improves, driving prices of properties up in areas where nature availability is high and vice versa. These dynamics are not accounted for in policy simulations based on second-stage hedonic studies, which rest on the assumption that the market is in equilibrium. This should be kept in mind, particularly when investigating larger policy changes that may affect market equilibrium. Klaiber and Phaneuf (2010) simulate policy scenarios using their sorting model to predict the policy response of prices. They find that the general equilibrium effects can extend well beyond the area directly affected by a policy. Walsh (2007) studies green space policies in a general equilibrium setting, allowing households to adjust private open space in response to changes in public open space supplied. He finds that such general equilibrium effects may be large and that increases in public open space can result in a reduction of overall open space available as households reduce the amount of land privately allocated to open space.

A number of important assumptions are made in our analyses, which merit further mention. This study is one of the few in the literature that moves beyond implicit prices in a hedonic study. Here, we apply a transparent identification strategy based on functional form restrictions, first suggested and applied by Bajari and Benkard (2005) and later applied by Bajari and Kahn (2005) and von Graevenitz (2018). To identify preferences, we impose an assumption about the functional form of utility, which can, of course, be criticized for being a strong restriction. However, the chosen form can be seen as a (good specific) local approximation for the unknown true form. The method is well suited for recovering and evaluating household-specific WTP for smaller changes that are non-marginal for the household, yet not affecting the market equilibrium. It is also suited for examining sources of preference heterogeneity as we do in this paper. Other approaches for obtaining identification, such as using an IV approach on multiple market models (e.g., Day et al., 2007), are sometimes less transparent and give rise to other theoretical and empirical obstacles that must be resolved.

The exact identification of a preference parameter relies on the household not choosing a corner solution, and thus we need the household to buy into the market to obtain more than just an upper bound for taste. Our analysis does not reveal much about the households with the lowest WTP for access to peri-urban nature areas.

A final comment relates to the use of the hedonic method for valuing these nature areas. This method implies a focus on use values to which you gain access by living nearby. However, people can hold non-use

Table A1
First stage results.

Spatial control	Fixed effect	Fixed effect and spatial smoothing	Fixed effect and clustered residual
Time control	Time smoothing	Time smoothing	Polynomial dates
Model	GAM (1)	GAM (2)	GLM (3)
Log (area)	0.44020***	0.43877***	0.44208***
	(0.02028)	(0.02028)	(0.03323)
Toilets	0.05011***	0.05091***	0.04796***
	(0.00905)	(0.00903)	(0.01074)
Garden	0.00013***	0.00014***	0.00013***
	(0.00001)	(0.00001)	(0.00002)
Roof: tile	0.05264***	0.05033***	0.05216***
	(0.00853)	(0.00859)	(0.00918)
Roof: Cement	0.04392**	0.04124**	0.04298**
	(0.01976)	(0.01975)	(0.02041)
Bathrooms	0.02663***	0.02632**	0.03013***
	(0.01028)	(0.01025)	(0.01129)
Rebuild in 70-ies	−0.02329*	−0.02162	−0.02579*
	(0.01387)	(0.01385)	(0.01420)
Rebuild in 00-s	0.07060***	0.07064***	0.07032***
	(0.02002)	(0.01999)	(0.02028)
Big roads within 400 m	−0.00016***	−0.00016***	−0.00015***
	(0.00004)	(0.00004)	(0.00005)
Nature within 800 m	0.00182***	0.00172***	0.00187***
	(0.00019)	(0.00019)	(0.00023)
Constant	3.86384***	3.76895***	3.86384***
	(0.09273)	(0.09196)	(0.14599)
AIC	28,312	28,204	28,312
Observations	2376	2376	2376
Adjusted R^2	0.544	0.582	0.544
Log Likelihood	−14,124	−14,102	−14,124

values wherever they live and people further away may also hold use values for these areas. We do not capture these values in our study.

6. Conclusion

Within the hedonic literature, very few studies move beyond implicit prices and onto estimating full demand schedules. This serves as a very thin basis for discussing the distributional impacts of public policy and for assessing preference heterogeneity. It may be the reason for the limited contribution from the hedonic literature to the discussion of distributional impacts from planning policies, such as urban greening policies, which we mainly find in the stated preference literature. The identification in the 2nd stage comes at the cost of a key assumption, which is a restriction on the functional form of the utility function. When evaluating smaller changes, yet large enough to be non-marginal to the individual, the restriction should allow for better results than policy analysis using implicit prices. However, for larger changes potentially affecting market equilibrium, the impact of such an identification strategy should be evaluated carefully. Recent contributions from the literature on sorting are better able to shed light on distributional aspects. However, these often require assumptions about preference heterogeneity for identification. Our study is complementary to these studies in examining sources and extent of preference heterogeneity in this context.

We use quantile regression to analyze whether the relationship between demographics and WTP is constant across the WTP distribution. We find WTP for peri-urban nature to increase across the entire WTP distribution with factors such as wealth and income. In contrast, higher education affects the center of the WTP distribution quite a bit but affects the tails to a lesser degree. The single and single-parent status also impacts WTP as does ownership of a car.

The detailed insight into the socio-demographic drivers of variation in preferences and hence WTP for the amenity values of peri-urban nature areas available around their homes has clear relevance for urban planning policies.

CRediT authorship contribution statement

Cathrine Ulla Jensen: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft, Writing - review & editing, Visualization. **Toke Emil Panduro:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration. **Thomas Hedemark Lundhede:** Methodology, Writing - review & editing, Supervision, Project administration. **Kathrine Graevenitz:** Conceptualization, Methodology, Writing - review & editing. **Bo Jellesmark Thorsen:** Conceptualization, Methodology, Writing - review & editing.

Table A2
Variation in wTP explained by demographics.

	OLS	Quantile regression 20%	40%	60%	80%
Intercept	4.274***	2.529***	4.183***	4.915***	6.116***
	(0.223)	(0.415)	(0.165)	(0.137)	(0.224)
Income (1000 EUR)	0.009***	0.008***	0.008***	0.008***	0.007***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Wealth (1000 EUR)	0.001***	0.001***	0.0004**	0.0004***	0.0005***
	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.00005)
Long education	0.259**	0.269*	0.508***	0.571***	0.278***
	(0.112)	(0.138)	(0.067)	(0.074)	(0.096)
Outside the workforce	−0.461	−1.365***	−0.824	−0.087	−0.228
	(0.377)	(0.494)	(0.913)	(0.348)	(0.317)
Self-employed	0.482***	0.855***	0.396***	0.306***	0.405***
	(0.151)	(0.216)	(0.135)	(0.101)	(0.126)
Topmanager	0.893***	1.181***	1.085***	0.816***	0.762***
	(0.149)	(0.235)	(0.143)	(0.098)	(0.107)
Car owner	−0.319	0.068	−0.346***	−0.349***	−0.818***
	(0.217)	(0.387)	(0.119)	(0.127)	(0.219)
Oldest member min 61	0.812***	1.207***	0.658***	0.532***	0.426***
	(0.099)	(0.136)	(0.063)	(0.049)	(0.046)
Single	0.617*	1.693***	0.478***	0.048	−0.284
	(0.359)	(0.370)	(0.141)	(0.162)	(0.946)
Single parent	−0.479	−1.154***	−0.245*	−0.083	−0.016
	(0.347)	(0.347)	(0.148)	(0.153)	(0.941)
R^2/Pseudo R	0.15	0.10	0.15	0.16	0.14

Note: N = 1864. *p < 0.1; **p < 0.05; ***p < 0.01

Table A3
First stage results with extremed priced housing.

Spatial control	Fixed effect	Fixed effect and spatial smoothing	Fixed effect and clustered residual
Time control	Time smoothing	Time smoothing	Polynomial dates
Model	GAM	GAM	GLM
Log (area)	0.46924***	0.46747***	0.46899***
	(0.02175)	(0.02175)	(0.03322)
Toilets	0.04977***	0.05018***	0.04787***
	(0.00979)	(0.00978)	(0.00975)
Garden	0.00016***	0.00017***	0.00016***
	(0.00001)	(0.00001)	(0.00003)
Roof: tile	0.05455***	0.05223***	0.05375***
	(0.00927)	(0.00932)	(0.01035)
Roof: Cement	0.06064***	0.05785***	0.05950***
	(0.02114)	(0.02114)	(0.01995)
Bathrooms	0.02855**	0.02871***	0.03154***
	(0.01109)	(0.01109)	(0.01112)
Rebuild in 70-ies	−0.04399***	−0.04261***	−0.04655***
	(0.01505)	(0.01504)	(0.01432)
Rebuild in 00-s	0.07501***	0.07599***	0.07347***
	(0.02127)	(0.02126)	(0.02074)
Big roads within 400 m	−0.00022***	−0.00021***	−0.00022***
	(0.00005)	(0.00005)	(0.00006)
Nature within 800 m	0.00220***	0.00219***	0.00224***
	(0.00019)	(0.00020)	(0.00025)
Constant	3.53405***	3.58519***	3.70526***
	(0.09536)	(0.09705)	(0.14851)
AIC	29,881	29,879	29,938
Observations	2448	2448	2448
Adjusted R^2	0.616	0.618	0.616
Log Likelihood	−14,941	−14,939	−14,937

Table A4
Summary statistics of the fixed effect school districts.

School district	N
Bagsværd skole	198
Buddinge skole	36
Engelsborgskolen	174
Fuglsanggårdsskolen	157
Gladsaxe skole	185
Hummeltofteskolen	206
Høje Gladsaxe skole	134
Kongevejens Skole	218
Lindegårdsskolen	88
Lundtofte Skole	106
Marielyst skole	47
Stengård skole	104
Søborg skole	68
Trongårdsskolen	174
Vadgård skole	228
Virum Skole	192
Værebro skole	61

Note: 2376 observations.

Appendices

References

Bajari, P., & Benkard, C. L. (2005). Demand estimation with heterogeneous consumers and unobserved product characteristics: a hedonic approach. *Journal of Political Economy, 113*(6), 1239–1276. https://doi.org/10.1086/498586.

Bajari, P., & Kahn, M. E. (2005). Estimating housing demand with an application to explaining racial segregation in cities. *Journal of Business & Economic Statistics, 23*(1), 20–33. https://doi.org/10.1198/073500104000000334.

Bartik, T. J. (1988). Measuring the benefits of amenity improvements in Hedonic P. *Land Economics, 64*(2), 172. https://search.proquest.com/docview/206727958?accountid=14468.

Belluzzo, W., Jr (2004). Semiparametric approaches to welfare evaluations in binary response models. *Journal of Business & Economic Statistics, 22*(3), 322–330. https://doi.org/10.1198/073500104000000172.

Bishop, K. C., & Murphy, A. D. (2011). Estimating the willingness to pay to avoid violent crime: a dynamic approach. *American Economic Review, 101*(3), 625–629. https://doi.org/10.1257/aer.101.3.625.

Brasington, D. M., & Hite, D. (2005). Demand for environmental quality: a spatial hedonic analysis. *Regional Science and Urban Economics, 35*(1), 57–82. https://doi.org/10.1016/j.regsciurbeco.2003.09.001.

Cajias, M., Fuerst, F., & Bienert, S. (2016). Are energy efficiency ratings ignored in the German housing market?–evidence from a large-sample hedonic study. Evidence from a large-sample hedonic study (June 22, 2016). SSRN Electronic Journal, June. https://doi.org/10.2139/ssrn.2799206.

Chattopadhyay, S. (1999). Estimating the demand for air quality: new evidence based on the Chicago housing market. *Land Economics, 75*(1), 22. https://doi.org/10.2307/3146991.

Cho, S.-H., Jung, S., & Kim, S. G. (2009). Valuation of spatial configurations and forest types in the Southern Appalachian Highlands. *Environmental Management, 43*(4), 628–644. https://doi.org/10.1007/s00267-008-9209-0.

Danish Business Authority. (2011). The Central Company Register (CVR). Danish Business Authority. https://datacvr.virk.dk/data/.

Day, B., Bateman, I., & Lake, I. (2007). Beyond implicit prices: recovering theoretically consistent and transferable values for noise avoidance from a hedonic property price model. *Environmental and Resource Economics, 37*(1), 211–232. https://doi.org/10.1007/s10640-007-9121-8.

De Valck, J., Landuyt, D., Broekx, S., Liekens, I., De Nocker, L., & Vranken, L. (2017). Outdoor recreation in various landscapes: which site characteristics really matter? *Land Use Policy, 65*, 186–197. https://doi.org/10.1016/j.landusepol.2017.04.009.

Diamond, R., & McQuade, T. (2019). Who wants affordable housing in their backyard? An equilibrium analysis of low-income property development. *Journal of Political Economy, 127*(3), 1063–1117. https://doi.org/10.3386/w22204.

ESRI 2011.. (2015). *ArcGIS Desktop*. Redlands, CA: Environmental Systems Research Institute.

Fritsch, M., Haupt, H., & Ng, P. T. (2016). Urban house price surfaces near a World Heritage Site: Modeling conditional price and spatial heterogeneity. *Regional Science and Urban Economics, 60*, 260–275. https://doi.org/10.1016/j.regsciurbeco.2016.07.011.

Jensen, F.S. (2003).Friluftsliv i 592 skove og andre naturområder. Skovbrugsserien nr. 32, Skov & Landskab, Hørsholm. https://core.ac.uk/download/pdf/269190408.pdf.

Jiao, L., & Liu, Y. (2010). Geographic field model based hedonic valuation of urban open spaces in Wuhan, China. *Landscape and Urban Planning, 98*(1), 47–55. https://doi.org/10.1016/j.landurbplan.2010.07.009.

Klaiber, H. A., & Phaneuf, D. J. (2010). Valuing open space in a residential sorting model of the Twin Cities. *Journal of Environmental Economics and Management, 60*(2), 57–77. https://doi.org/10.1016/j.jeem.2010.05.002.

Koenker, R. (2013). Quantreg: quantile regression. R Package Version 5.

Koenker, R., & Hallock, K. (2001). Quantile regression: an introduction. *Journal of Economic Perspectives, 15*(4), 43–56. https://doi.org/10.1257/jep.15.4.43.

Kong, F., Yin, H., & Nakagoshi, N. (2007). Using GIS and landscape metrics in the hedonic price modeling of the amenity value of urban green space: a case study in Jinan City, China. *Landscape and Urban Planning, 79*(3-4), 240–252. https://doi.org/10.1016/j.landurbplan.2006.02.013.

Lake, I. R., Lovett, A. A., Bateman, I. J., & Day, B. (2000). Using GIS and large-scale digital data to implement hedonic pricing studies. *International Journal of Geographical Information Science, 14*(6), 521–541. https://doi.org/10.1080/136588100415729.

Li, Chuan-Zhong, & Mattsson, Leif (1995). Discrete choice under preference uncertainty: an improved structural model for contingent valuation. *Journal of Environmental Economics and Management, 28*(2), 256–269. https://doi.org/10.1006/jeem.1995.1017. In this issue.

Mansfield, C., Pattanayak, S. K., McDow, W., McDonald, R., & Halpin, P. (2005). Shades of green: measuring the value of urban forests in the housing market. *Journal of Forest Economics, 11*(3), 177–199. https://doi.org/10.1016/j.jfe.2005.08.002.

Morancho, A. B. (2003). A hedonic valuation of urban green areas. *Landscape and Urban Planning, 66*(1), 35–41. https://doi.org/10.1016/S0169-2046(03)00093-8.

Netusil, N. R., Chattopadhyay, S., & Kovacs, K. F. (2010). Estimating the demand for tree canopy: a second-stage hedonic price analysis in Portland, Oregon. *Land Economics, 86*(2), 281–293. https://doi.org/10.3368/le.86.2.281.

Notaro, S., & De Salvo, M. (2010). Estimating the economic benefits of the landscape function of ornamental trees in a sub-Mediterranean area. *Urban Forestry & Urban Greening, 9*(2), 71–81. https://doi.org/10.1016/j.ufug.2009.09.001.

O'Garra, T., & Mourato, S. (2007). Public preferences for hydrogen buses: comparing interval data, OLS and quantile regression approaches. *Environmental and Resource Economics, 36*(4), 389–411. https://doi.org/10.1007/s10640-006-9024-0.

Orford, S. (2002). Valuing locational externalities: a GIS and multilevel modelling approach. *Environment and Planning B: Planning and Design, 29*(1), 105–127. https://doi.org/10.1068/b2780.

Palmquist, R. B. (1992). Valuing localized externalities. *Journal of Urban Economics, 31*(1), 59–68. https://doi.org/10.1016/0094-1190(92)90032-G.

Panduro, T. E., Jensen, C. U., Lundhede, T. H., von Graevenitz, K., & Thorsen, B. J. (2018). Eliciting preferences for urban parks. *Regional Science and Urban Economics, 73*, 127–142. https://doi.org/10.1016/j.regsciurbeco.2018.09.001.

Panduro, T. E., & Veie, K. L. (2013). Classification and valuation of urban green spaces—a hedonic house price valuation. *Landscape and Urban Planning, 120*, 119–128. https://doi.org/10.1016/j.landurbplan.2013.08.009.

Phaneuf, D. J., & Requate, T. (2016). *A course in environmental economics: theory, policy, and practice*. Cambridge University Press.
Poudyal, N. C., Hodges, D. G., & Merrett, C. D. (2009). A hedonic analysis of the demand for and benefits of urban recreation parks. *Land Use Policy, 26*(4), 975–983. https://doi.org/10.1016/j.landusepol.2008.11.008.
R Core Team. (2015). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy, 82*(1), 34–55. https://doi.org/10.1086/260169.
Sander, H., Polasky, S., & Haight, R. G. (2010). The value of urban tree cover: a hedonic property price model in Ramsey and Dakota Counties, Minnesota, USA. *Ecological Economics, 69*(8), 1646–1656. https://doi.org/10.1016/j.ecolecon.2010.03.011.
SKAT. (2012). Den offentlige informations server. https://www.ois.dk/.
Ta, M. T., Tardieu, L., & Levrel, H. (2020). Specifying preference heterogeneity regarding natural attributes of urban green spaces to inform renaturation policies (No. 2020-78). CIRED Working Paper.
The Danish Geodata Agency. (2011). kort10. Vol. 2011. http://download.kortforsyningen.dk/.
Tu, G., Abildtrup, J., & Garcia, S. (2016). Preferences for urban green spaces and peri-urban forests: an analysis of stated residential choices. *Landscape and Urban Planning, 148*, 120–131. https://doi.org/10.1016/j.landurbplan.2015.12.013.
Tuffery, L. (2019). Revealed preferences for forest recreational services using the random bidding model. *Landscape and Urban Planning, 189*, 181–188. https://doi.org/10.1016/j.landurbplan.2019.04.022.
Tyrväinen, L., & Miettinen, A. (2000). Property prices and urban forest amenities. *Journal of Environmental Economics and Management, 39*(2), 205–223. https://doi.org/10.1006/jeem.1999.1097.
Tyrväinen, L., Mäkinen, K., & Schipperijn, J. (2007). Tools for mapping social values of urban woodlands and other green areas. *Landscape and Urban Planning, 79*(1), 5–19. https://doi.org/10.1016/j.landurbplan.2006.03.003.
van den Berg, A. E., Koole, S. L., & van der Wulp, N. Y. (2003). Environmental preference and restoration: (how) are they related? *Journal of Environmental Psychology, 23*(2), 135–146. https://doi.org/10.1016/S0272-4944(02)00111-1.
Vining, J., Merrick, M. S., & Price, E. A. (2008). The distinction between humans and nature: human perceptions of connectedness to nature and elements of the natural and unnatural. *Human Ecology Review, 1–11*. https://doi.org/10.2307/24707479.
von Graevenitz, K. (2018). The amenity cost of road noise. *Journal of Environmental Economics and Management, 90*, 1–22. https://doi.org/10.1016/j.jeem.2018.04.006.
von Graevenitz, K., & Panduro, T. E. (2015). An alternative to the standard spatial econometric approaches in hedonic house price models. *Land Economics, 91*(2), 386–409. https://doi.org/10.3368/le.91.2.386.
Walsh, R. (2007). Endogenous open space amenities in a locational equilibrium. *Journal of Urban Economics, 61*(2), 319–344. https://doi.org/10.1016/j.jue.2006.09.002.
Willis, K., & Crabtree, B. (2011). Measuring health benefits of green space in economic terms. In K. Nilsson, M. Sangster, C. Gallis, T. Hartig, S. de Vries, K. Seeland, & J. Schipperijn (Eds.), *Forests, trees and human health* (pp. 375–402). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-90-481-9806-1_13.
Wood, S. N. (2017). *Generalized additive models: an introduction with R*. CRC Press.