CIRED Working Paper
N° 2020-78 - Février 2020

Specifying preference heterogeneity regarding natural attributes of urban green spaces to inform renaturation policies

Mai-Thi Ta
Léa Tardieu
Harold Levrel

Corresponding author: maithi.ta@centre-cired.fr

Centre international de recherche sur l'environnement et le développement
Unité mixte de recherche CNRS - Ecole des Ponts ParisTech - Cirad - EHESS - AgroParisTech
Site web: www.centre-cired.fr Twitter: @cired8568
Jardin Tropical de Paris -- 45bis, avenue de la Belle Gabrielle 94736 Nogent-sur-Marne Cedex

Specifying preference heterogeneity regarding natural attributes of urban green spaces to inform renaturation policies

Abstract

It is now acknowledged that greening cities influences positively the quality of life of dwellers through several ways, such as increasing outdoor recreational opportunities, improving in turn social connections, physical or mental health. However, in urban planning, it is usually implicitly assumed that any type of urban green space will have the same effect on citizen's well-being. The location and design of urban greening investments are barely compared to the needs of the city dwellers. This paper addresses the demand side of urban greening by questioning which green spaces' characteristics people value. To do so, we applied a distance-based choice experiment to capture the trade- off between the green spaces' attributes (e.g. tree cover, size, shape) and the travel time citizens' are willing to spend to reach a hypothetical site. We found a strong heterogeneity among the interviewees' expectations regarding a green space. We show that dwellers' tastes are influenced by many parameters such as geographical and socio-economic characteristics, or the activities they are used to practice. The results suggest that accounting for potential users' preferences regarding urban nature characteristics is critical to inform urban planning in order to build efficient designs of urban greening programs.

Keywords: Green spaces; recreational services; renaturation policies; preference heterogeneity, choice experiment; urban nature

1. Introduction

The United Nations recently set an objective of "[providing] by 2030 a universal access to safe, inclusive and accessible, green and public spaces" less than 300 meters from each inhabitant's residence (Sustainable Goal 11.7, United Nations Department of Economic and Social Affairs, 2014). This recommendation was formulated based on the growing recognition of the positive influence of urban nature[1] on the quality of life of dwellers in terms of social connections (Luck and al., 2011), physical and mental health (Alcock and al., 2014; Coppel and Wüstemann, 2017), children's cognitive development (Markevych and al.,10 2014), and reduction of anxiety (Cox et al., 2017).

In the Paris metropolitan region (Ile-de-France), the UN 11.7 goal is far from being achieved, as only 66% of residents have access to a public urban green space of at least 1.5 ha less than 300 meters from their residence (Tardieu et al. in progress). This percentage falls to 50% in the Paris inner city, making Paris one of the least endowed urban areas in terms of green spaces in France. Moreover, Paris is classified as one of the cities with the lowest green view indexes in the world[2]. Even though access to green spaces increased by 2% from 1982 to 2017 in the region and in Paris inner city, the surface area of urban green space per inhabitant decreased by 24% during this period due to net conversion of natural and open areas of roughly 40 700 hectares (Tardieu et al., in progress). This artificialization trend goes against public opinion, as 53% of Parisians consider that insufficient land is allocated to nature[3]. Therefore, to meet the United Nation's objective, increased control of land artificialization and investment in renaturation policies are needed.

The Ile-de-France region is attempting to follow the United Nation's recommendation. For instance, in 2013, the region fixed a regulatory objective of access to be achieved by 2030 in the main land planning document: the SDRIF[4] (regional plan for Ile-de-France region). The SDRIF aims to supply roughly 10 square meters of open area per inhabitant in the region and gives priority to municipalities with less than 10% of open and natural areas.

However, although an important step was taken by including this objective in the SDRIF, the local population's preferences regarding urban nature are ignored, potentially making this policy inefficient. Indeed, inhabitants may show heterogeneous preferences regarding different attributes of accessible urban natural areas. To ensure the effectiveness of the policy, decision makers need spatial information about residents' preferences. Investing in renaturation alone will not ensure that neighboring residents actually benefit from these areas. Urban planners therefore need to better understand residents' use of green spaces and its drivers.

In an attempt to provide some answers to these questions, we conducted a discrete choice experiment (DCE) in Ile-de-France to specify residents' expectations regarding urban green areas. DCE was first applied in environmental economics in the late 90s (Adamowicz et al., 1994; Hanley et al., 1998) after having been developed as a marketing

[1] Our use of "urban green spaces" encompasses everything from small patches such as bioswales and green roofs to large urban parks and peri-urban forests.

[2] According to The Green View Index developed by the Massachusetts Institute of Technology calculated using Google Street View panoramas, showing the percentage of canopy coverage of a particular location: http://senseable.mit.edu/treepedia/cities/paris

[3] https://www.ifop.com/wp-content/uploads/2019/05/116361-Rapport-Le-retour-à-la-campagne.pdf

[4] The SDRIF- Shéma Directeur Régional de la région Ile-de-France- sets the land management policies for the region. The last SDRIF approved in 2013 contains plans for overall urban development at the regional scale through 2030. The plan selects areas to densify, urban lands to extend, transport infrastructure to develop, and green areas to enhance, among other goals.

tool to evaluate consumer behavior (Green & Rao, 1971; Green et al., 1972). Since then, DCE has increasingly been used to characterize consumer preferences regarding various types of goods and services: e.g. agri-environmental contracts for farmers (Espinosa-Goded et al., 2010; Vaissière et al., 2018), recreational demand for natural areas (De Valck et al., 2017; Mulatu, 2019; Tu et al., 2016), perception of biological invasions (Chakir et al., 2016) and mitigation policies (Kermagoret et al., 2016).

Despite recent DCE development in environmental economics, few case studies have focused on urban nature. De Valck et al. (2017) studied preferences for outdoor recreational destinations in Antwerp province (Belgium). Tu et al. (2016) investigated preferences regarding the distance to urban green spaces from respondent's houses in Nancy (France). Mulatu (2019) explored preferences for urban green spaces in Addis Ababa, Ethiopia.

Among those studies, two found preference heterogeneity among respondents and attempted to identify its factor using additional information on respondents' socio-economic characteristics and hobbies. De Valck et al. (2017) showed that preferences for outdoor recreation vary according to the outdoor activities that respondents' practice. Tu et al. (2016) related the heterogeneity of residential choice preferences to the frequency of visits to a forest and the access of residents' houses to private gardens.

This kind of data is not readily available to decision makers, who generally have easier access to general socio-economic data such as age, income or socio- professional categories. In the studies we identified, only Tu et al. (2016) relate preferences heterogeneity to easily accessible data: they show that willingness-to-pay to live close to a park decreases with income.

Our study reinforces the literature on DCE applied to urban ecosystems and make two novel contributions. First, the study is action-oriented, as we use "distance-based" DCE and a latent class model. Limited research has been conducted regarding individuals' decision processes leading to the choice of a recreational destination, especially the trade-off between the distance to the recreational site and its characteristics (de Valck et al., 2017). The distance-based DCE approach measures the willingness to spend time travelling to a recreational green space according to its different attributes rather than willingness-to-pay. We believe that evaluating willingness to spend time rather than willingness to pay is more engaging as the value of time is greater than the value of money for many respondents. Second, we identify three different profile of individuals with differing preferences using a latent class model. Policy makers can use these results to spatially adjust their actions according to the socio-economic characteristics of a neighborhood.

This paper is structured in 3 main sections. Section 2 details the choice experiment and econometric methodologies and describes the materials used for the survey design. Section 3 presents the results of the survey. Finally, section 4 synthesizes our major findings and explains how they could be used in a decisional context.

2. Material and Methods

2.1 Case study

Our study focuses on the French region of Ile-de-France. The region is com- posed of 8 departments, which are divided into Paris ***"intra-muros",*** the ***"little crown"*** - inner suburbs of Paris containing Hauts-de-Seine, Val de Marne and Seine-Saint-Denis, and the ***"big crown"***- outer suburbs containing Val-d'Oise, Yvelines, Essonne and Seine-et-Marne.

Despite representing only 2% of the French territory, this region accounts for 18% of the population and generates 31% of the national growth domestic product. The region attracts many qualified workers, as 35% of French executives live in the region. Moreover, it is a young region, with a mean age of 37.8 years (vs 40.9 at the national level).

Ile-de-France remains the most economically unequal French region: the overall poverty rate is 16% (one point higher than the national rate) and peaks at 38% in some northern cities. Five of its departments figure among the richest in the country, while Seine St-Denis is among the poorest. Inequalities continue to grow as the highest revenues increase more rapidly than the poorest. In Paris, the gap is even larger: the standard of living of the 10%-poorest households is lower than the regional level[5].

The Ile-de-France region is covered by mostly agricultural areas, which occupy nearly 50% of the territory, situated mainly in Paris's big crown, followed by woods and forests (Figure 1). The region benefits from 8342 km of water- ways, 30 000 ponds, and three national parks (Haute Vallée de Chevreuse, Vexin and Gatinais).

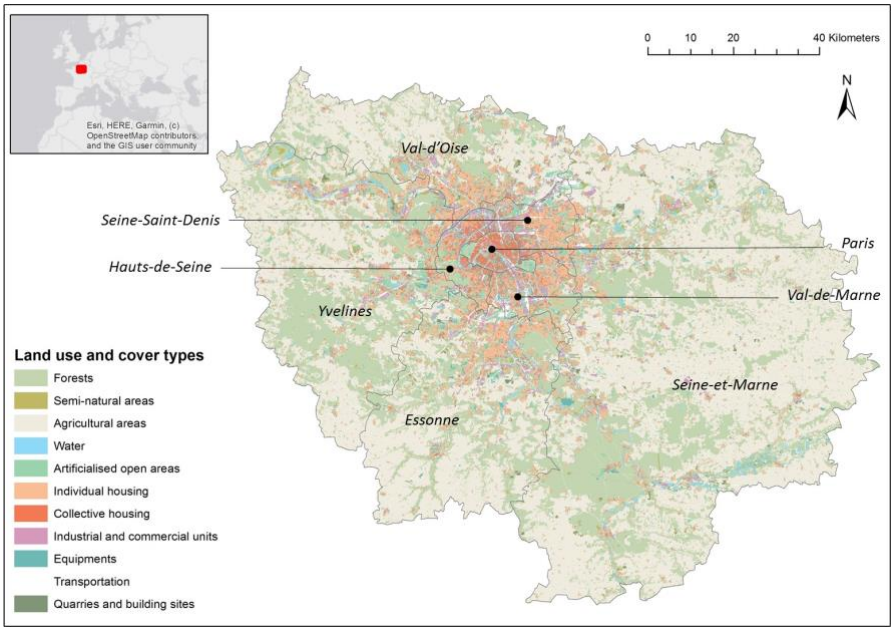


Figure 1: Land use and land cover in the Ile-de-France region, France (based on the Mode d'Occupation des Sols developed by the Institut d'Aménagement Urbain)

Globally, since 2012, the surface area of green spaces increased after a long period of decrease (Tardieu et al., in progress). The core of the metropolis also experienced a positive dynamic, with an average increase in the area of urban open spaces of 13 ha/year. Open spaces refer to parks and gardens (whether private or not, excluding castle gardens), outdoor sports fields, golf courses, leisure parks, and grasslands. Although, it is too early to assess the precise role of SDRIF compared to that of the decrease in economic activity, the situation of the most deprived municipalities, remained unchanged without deteriorating, despite the regulatory guidelines aimed at re-balancing the supply of green spaces. Figure 2 shows a map indicating the share of a municipality's population living 300 meters or less from a green space of at least 1.5 ha. In the "big crown" residents have a large access to green spaces while access is more limited in the ***"little crown"*** and weak

[5] Insee Flash Ile-de-France no39, Entre 2013 et 2015, les écarts de revenus se sont creusés entre départements franciliens https://www.insee.fr/fr/statistiques/3717097

in Paris ***"intra-muros"***.

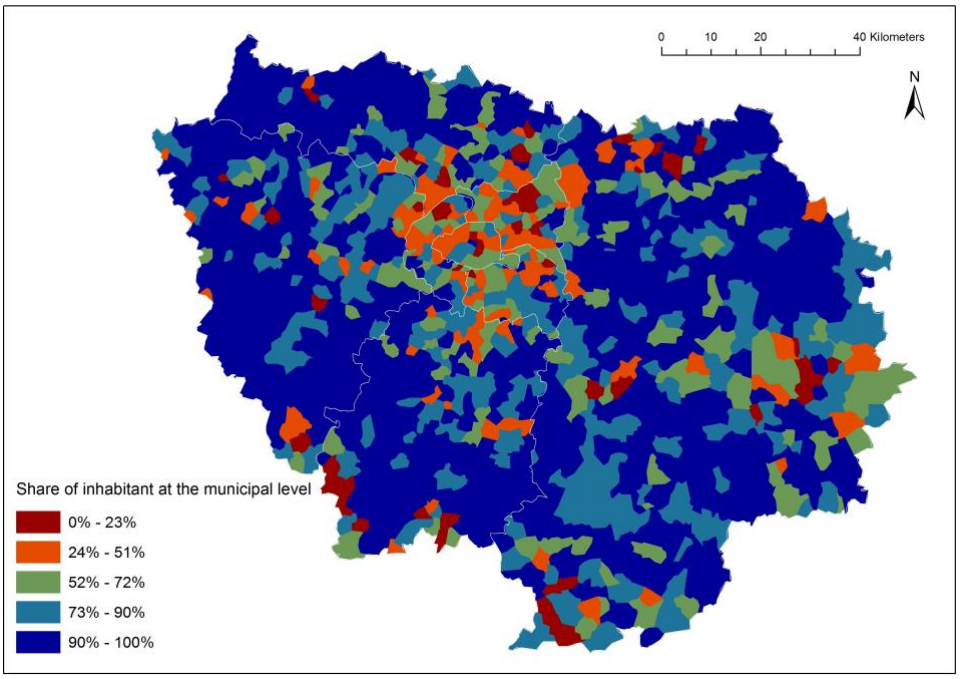


Figure 2: Share of inhabitants at the municipal level having access in 2017 to a recreational green space

2.2 Methodological options

We rely on DCE to specify the preferences of Ile-de-France inhabitants for outdoor recreation activities related to green spaces. DCEs are based on micro-economic theory of consumer choice and non-market valuation techniques. Respondents are asked to choose between different options of a good or service, described in terms of its components. One component is a monetary variable that is used to estimate the marginal rate of substitution. A choice situation is generally composed of two options and a status quo (i.e., the current situation where nothing is changed). Each respondent goes through several choice situations corresponding to the choice experiment. To obtain a relatively short and easily understandable survey, the number of attributes describing the good or service is limited to the most representative.

2.2.1 Modelling framework

Lancaster's theory and random utility theory. The choice experiment modelling framework is based on Lancaster's theory of value (1966) and random utility theory (McFadden et al., 1973). Lancaster's theory (1966) argues that consumers perceive a good or a service through its different characteristics and that consumers can distinguish their preferences for each characteristic. For ex- ample, the value of a computer consists of the value of its screen size and weight plus the value of its processor, memory card, etc. In a choice experiment, the value of each alternative corresponds to the sum of the value given to each of its attributes' levels. Respondents select one of the two alternatives by comparing their utility derived from each alternative.

McFadden et al. (1973) consider that some randomness is involved when people make a decision: the utility function contains a deterministic component and a random component. The deterministic part is composed of observable individual characteristics, whereas the random part accounts for unobserved variables that influence the choice. The random utility $U_{n,i}$ that an individual *n* obtains from alternative i is:

$$U_{i,n} = V_{i,n} + \epsilon_{i,n} \qquad (1)$$

where $V_{i,n}$ is the deterministic component and $\epsilon_{i,n}$ is the stochastic element.

A respondent *n* will choose alternative i over any other alternative *j* in a choice situation if the utility he obtains from alternative i is larger:

$$U_{n,i} > U_{n,j} \qquad \forall j \neq i$$

2.2.2. Econometric specification

We use different econometric models to analyse DCEs: the conditional logit, mixed logit and latent class models.

The conditional logit model. The conditional logit model assumes that the error terms are independent and identically distributed (IID). Among all agents and alternatives, the random component follows a standard Gumbel distribution. The conditional logit model also assumes independence from irrelevant alternatives (IIA): the preference for one alternative is independent of that for any other alternative. This implies that the unobserved components of utility are uncorrelated between two alternatives.

Let $X_{n,i}$ be the attribute levels of alternative i chosen by individual n; we postulate that utility is linear. The probability that individual n chooses alternative i over any other alternative ***j*** can be written as:

$$\mathbb{P}_{n,i} = \frac{\exp(\beta' X_{n,i})}{\sum_j \exp(\beta' X_{n,j})} \qquad (2)$$

where β' is a vector of the parameters estimated via maximum likelihood.

Appendix C provides more details about the model's computations.

Caveats of conditional logit models. The conditional logit model assumes that preferences are the same among all individuals, when in reality, we respondents might have unique personal preferences. The model also assumes that the errors are not correlated between respondents and therefore does not allow any variable to influence the unobserved part of the utility. However, we could assume that unobserved utilities are influenced by a variable between respondents and are thus correlated. A random parameter logit (or mixed logit) can be implemented to include this possible correlation in preferences and to account for heterogeneity among respondents.

Mixed logit. The mixed logit model is free of the IIA hypothesis and allows for heterogeneity in preferences. The model estimates the parameters of the assumed distribution of

preferences among the population using a simulated maximum likelihood. In each step of the optimization, R draws of the parameters β are taken from the assumed distribution, and the corresponding likelihood is computed. The values that maximize the simulated likelihood are parameter estimates (mean and standard deviation for a normal distribution). More details of the intermediary computations are given in Appendix B.

Willingness to accept. The willingness to accept (WTA) a unit change in an attribute corresponds to its marginal rate of substitution with the monetary attribute. In our case, the monetary attribute is not properly expressed in monetary terms as it relies on the time spent to travel to a green space. The WTA is calculated as follows:

$$WTA_{X_k} = -\frac{\beta_{X_k}}{\beta_{time}} \tag{3}$$

where β_{Xk} is the estimated coefficient for the considered attribute X_k.

Latent class. Latent class models are similar to mixed logit models but are free from assumptions on the distribution of preferences among respondents. The population is assumed to be divided into Q classes. Individuals in a given class exhibit homogeneous preferences, but these preferences differ from those of individuals in other classes.

A priori, we hypothesize that no socio-economic variable determines class membership. We thus have a constant class membership among individuals, which allows us to compute the unconditional probability of the sequence of choices made by individual *n* (see Appendix C for detailed steps of the method).

Individual-specific estimates. The discrete distributions of the parameters give an overview of how preferences are distributed in the population. We calculate individual-specific parameters to determine where an individual is situated in the distribution given a set of choices. We estimate the posterior membership probability that individual n belongs to class *q* given the observed choices (Greene and Hensher, 2003; Kamakura and Russell, 1989). The expression is given in equation C.3.

Relating class membership to auxiliary variables. The latent class model identifies groups of individuals presenting homogeneous preferences. One could explore what individual characteristics influences class membership. To do that, we regress individual's logit-transformed posterior probability, defined as π_{nq} in equation A.5, on auxiliary variables (Clarck & Muthén, 2009; Vermunt, 2010). Formally, the logit regression specification is:

$$\log\left(\frac{\hat{\pi}_{nq}}{1-\hat{\pi}_{nq}}\right) = \beta_1 X_1 + \ldots + \beta_K X_K \tag{4}$$

where YC defines the membership of individual n to class q and $X_k = (x1,\ldots,xK)$ is the set of auxiliary variables. To evaluate the influence of a variable X_k, we calculate the odds ratio when the variable increases by 1 unit:

$$\text{Odds ratio} = \frac{Odds_{X_k+1}}{Odds_{X_k}} = \exp(\beta_k) \tag{5}$$

Thus, if $\exp(\beta_k) > 1$, X_k has a positive influence on the membership probability of class Q.

2.3 Survey and questionnaire description

2.3.1 Questionnaire development

The questionnaire contains three main parts (Appendix F). The questionnaire began with a short introduction explaining the aim of the study. Then, the first part of the survey consisted of questions regarding respondent's socio- economic situation (place of residence, sex, socio-professional category, house- hold composition). Next, the choice sets were presented, with careful attention given to provide neutral descriptions. The order in which the choice sets were presented was randomized to avoid having less attention dedicated to the same choice sets. The third part of the questionnaire asked follow-up questions regarding respondents' habits concerning green spaces, their revenue and diploma.

Selection of attributes and definition of attribute levels. To determine which attributes to include in our survey, we interacted with stakeholders of the IDEFESE project[6]: French ministry of the environment, decentralized state services, urban planners, local authorities, agencies, developers, associations and think tanks or academic institutions.

To maximize the response rate, we wanted the questionnaire to last a maximum of 10 min. Thus, we chose to limit the number of attributes to six. Those attributes had to allow respondents to picture a realistic green space while providing additional information to be able to relate preferences to individual- specific variables.

- The first attribute is ***forest cover***, which represents the degree of wilder- ness, taking the density of tree cover as an indicator. The attribute takes three levels: landgrass, woodland or forest.
- The second attribute describes the ***shape*** of the fictional green space: either a linear form (as for a riverbank) or not.
- The third attribute specifies the presence (or not) of a ***water*** body in the fictional green space.
- The fourth attribute describes the ***size*** of the fictional green space. We follow commonly used differentiation: a green space is considered to be large if it is bigger than 1.5 ha (Cabral et al., 2016; Levrel et al., 2017; Niemelä et al., 2010). Otherwise, the green space is considered to be small.
- Because the Ile-de-France region is very dense, transportation mode considerably influences many everyday choices. The fifth attribute is thus the ***transport mode*** by which the green space is accessible. The attribute takes 3 levels: walking, biking, and taking public transport or a car. To estimate the extent to which biking can be used as an alternative transportation mode, we isolated biking as a unique level, whereas public transport and driving were combined as we considered them to be substitutes.

  If the respondent chose the "public transport or car", we added a follow- up question asking whether the respondent would prefer the car or public transport.
- The last attribute, ***transportation time***, indicates how long it takes to travel from the respondent's house to the green space. For this attribute, the lowest level is 5 min, which corresponds to a 300-meter walking trip (distance recommended by the UN in the AA.7 Sustainable goal). The highest level is 30 min, which is 5 min less than the

[6] See https://idefese.wordpress.com/ for more information about IDEFESE project.

average commuting time spent in the little crown and Paris to travel to work[7]. Between these limits, we added 2 more levels: 10 and 20 min.

Experimental design and choice experiment description. The second part of the questionnaire is the choice experiment. Choice cards were defined with the presentation of two hypothetical green spaces described by attributes taking different levels and a status quo option. Since we had 6 attributes, each taking between 2 and 4 levels, the full factorial design gave 296 combinations.

The presentation of such a design would have been cognitively too complicated for respondents. Therefore, we selected choice sets that would provide the most information about respondents' preferences. To do so, we first used SAS software to define the minimum number of choices sets to present to respondents. We then used N-Gene software, which gave a statistically optimal sub-set of the combinations of the full factorial design using a Bayesian D-optimal design. The 100% efficient design led to 72 or 144 choice sets, which was still too many for a respondent to answer. We chose to present 24 combinations with 1 violation. Gathered in pairs, this selection provided 12 choice scenarios. Figure 3 is an example of a choice set. The first column gives the attributes list, and the following two columns give the values for the two suggested options. The third column is the status quo alternative. We added an opt-out option because it is supposed to help immerse the respondent in the survey (Adamowicz and Boxall, 2001; Kontoleon and Yabe, 2003).

Table 1: Attributes and levels

Variables	Description	Levels	Coding Variable
Forest cover	Density of trees covering the green space	Grassland Wooded park Forest	Reference level arbore foret
Shape	Shape of the green area	Linear Not linear	lin
Water	Presence of water in the green space	Yes No	eau
Size	Size of the green space: either larger than 1.5 ha or not	Large Small	taille
	Type of transportation used to travel	By foot By bike	Reference level

Follow-up questions. The third part of the questionnaire is composed of several follow-up questions that aim to control different biases and to fill in other socio- economic characteristics of respondents.

First, a question regarding systematically omitted attributes was asked, which allowed us to identify protest answers (respondents who systematically ignored most of the attributes) and remove them from the analysis. Next, we asked a set of questions aimed at better characterizing respondents' profiles and helping us interpret the results. A large range of follow-up questions were asked about habits concerning green space attendance, ecological sensitivity and socio-economic profile. To identify different usage, we asked about the three main activities respondents participate in when visiting a green space. We also asked about the frequency of their visits to green spaces, their most frequently

[7] Les temps de déplacement entre domicile et travail, Dares Analyses, Publication from the Directorate for Research, Studies and Statistics (Direction de l'Animation de la Recherche, des Etudes et des Statistiques - DARES) , November 2015

visited green space, how long it takes on average to travel to a green area and how long they stay. We ended the questionnaire with two socio-economic questions about the individual's highest diploma and their current income.

Green space attributes	Alternative 1	Alternative 2	Status quo
Forest cover			I choose neither of both green spaces
Shape			
Presence of a water body			
Size	Large* / Long*	Small** / Short**	
Transport mode	Biking / Trotinette	By foot	
Transportation time	5 min	30 min	
Choice			2

Figure 3: Example of a choice set

2.4 Survey design

The surveys were delivered to 320 people from April 15th to May 24th 2018, representing 11 520 choices made. Given the short available time, we were assisted in the field by 5 students from Ecole des Ponts. We conducted privileged face-to-face interviews because online surveys have low response rates. Further- more, such interviews provided a good way to reach the less-connected portion of the population (elderly, poor) To avoid endogenous selection bias, careful attention was drawn to interviewing people outside of green spaces. To select the cities, we classified the municipalities of the region regarding 3 criteria:

- density: ratio between the municipality's population and its surface.
- median standard of living: ratio between household available income and consumption unit[8].
- urbanization ratio: ratio between a municipality's urbanized surface and its total surface. Urbanized and non-urbanized surface were determined based on the French land use classification in 11 posts[9]. Areas considered as non-urban surface were those classified as forest, semi-natural areas, farmed lands and water[10]. We classified open areas with some infrastructure (such as running trails, man-made

[8] INSEE calculated consumption unit based on OECD equivalence scale: the first household adult accounts for 1 consumption unit (UC), all other persons aged 14 years or older account for 0.5 UC each, and children younger than 14 years account for 0.3 UC each.

[9] The document is named Mode d'Occupation des Sols (MOS)

[10] Classified from 1 to 4 in the MOS year 2017 in 11 posts available at https://www.iau-idf.fr/fileadmin/DataStorage/IauEtVous/CartesEtDonnees/Mos/NomenclatureMOS-11-24-47-81.pdf

urban parks, or cemeteries) as non-urban areas (class 5 of the MOS). All other types of land-use[11] - mostly housing, lands dedicated to economic activities, equipment, trans- port, quarries and landfills - were considered to be urban surface.

Each criterion was divided into classes to cluster municipalities into groups. To determine the optimal cut, we used Jenks classification method (see appendix Appendix E). This method minimizes the variance within classes while maximizing the variance between classes. We fixed a number of 3 classes per criterion, which gave us 27 combinations, corresponding to 27 types of municipalities.

Those 27 combinations were theoretical: a corresponding municipality did not exist for all of them. After examining the distribution of municipalities, we concluded that only 17 combinations actually exist. For each unique combination, we randomly selected a municipality.

Table 2: Jenks classification results

	Low	Medium	High
Urbanization ratio	< 0.2166	[0.2166, 0.5627]	> 0.5627
Density (hab/km2)	< 3940.66	[3940.66; 14548.61]	> 14548.61
Standard of living (e/ UC))	< 9276	[9276; 19025]	> 19025

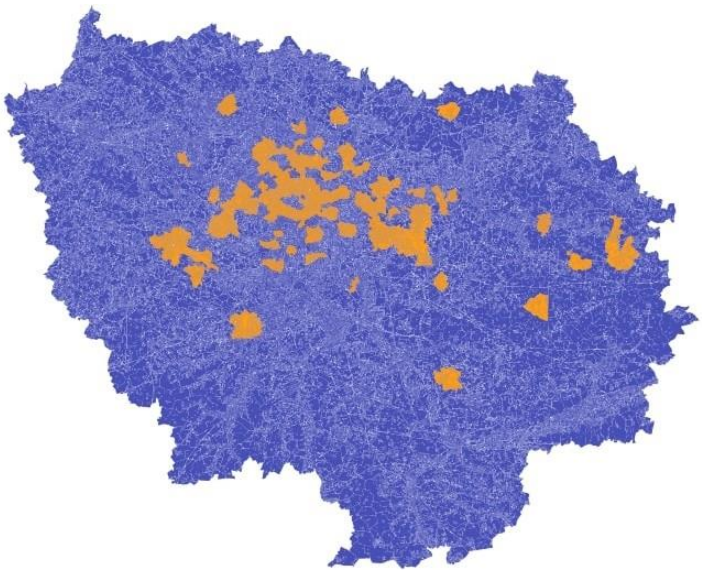


Figure 4: Geographical distribution of respondents

Despite targeting the 17 selected cities, it was difficult to avoid interviewing people from other municipalities. Figure 4 represents all cities in which at least one individual of the

[11] From number 6 to 11 of the French MOS, year 2017, 11 posts)

survey lives.

A total of 320 respondents filled in the questionnaire, but 13 did not complete all the choice situations, leaving 307 useful responses (corresponding to 11 052 observations). Table 3 gives the gender, age, socio-professional category, and revenue of the sample population. In terms of age and sex distribution, our sample is representative of the region: 52% of the respondents are women and two-thirds of the sample is less than 44 years old. The age distribution matches that of the region, although the 45 - 59 year-old group is slightly under-represented. Our sample under-represents intermediate professions and over-represents employees, possibly because we interviewed people in the streets during working hours, when most workers are not outside. We rectified this level of representativeness via weighting in the econometric models. Overall, our sample is small but sufficiently large to draw conclusions. Moreover, the sample has been corrected to be representative of the population.

Model specification. An alternative specific constant (ASC) is added, corresponding to the situation where the respondent chose neither of the two alternatives. The ASC is a dummy variable taking a value of 0 if one of the two fictional green spaces is chosen and a value of 1 if the opt-out option is chosen. The value reflects the utility if all attribute levels are null and captures the influence of unobserved variables. A significant ASC indicates the existence of omitted variables that influence the utility of respondents, leading them to prefer to not go to one of the suggested alternatives.

The estimated utility is:

$$U_{n,i} = \beta_0 ASC + \beta_{i,1} x_{i,1} + \ldots + \beta_{i,m} x_{i,m} + \ldots + \beta_{i,6} x_{i,6} + s_{n,i}$$

where $x_{i,m}$ with m $\in$ [1, 6] corresponds to the level taken by the attributes forest cover, shape, water, size, mode of transport and transport time.

Table 3: Sample and population characteristics

		Proportion of respondents in the sample (%	Proportion of the population in Ile-de-France (%
Gender	Male	48.0	48.0
	Female	52.0	52.0
Age class	[18,29]	30.1	No data
	[30,44]	31.4	
	[45,59]	18.6	
	[60,74]	15.7	
	Older than 75	4.2	
Socio-professional category	Farmer	0.1	0.1
	Craftsman, retailer, entrepreneur	6.3	3.2
	Managerial and higher-education professions	17.6	17.7
	Intermediate professions	8.5	16.1
	Employees	29.5	16.8
	Workers	4.4	8.7
	Retirees	14.2	19.7
	Other without professional activity	19.5	17.8
Monthly disposable revenue	< 1400e	32.5	No data
	1400eto 2000e	35.0	
	2000eto 3000e	20.0	
	> 3000e	12.5	

3. Results

3.1 Econometric models

3.1.1 Conditional logit

We first used a conditional logit to model respondents' choices. As shown in section 2.2.1, the conditional logit model assumes that the IIA hypothesis is relevant. The results should not be influenced by any additional alternative. If this hypothesis is not true, conditional logit models are not relevant. To test for the IIA hypothesis, we conduct the Hausman test.

The test compares the model coefficients $\boldsymbol{\beta_i}$ estimated on a subset of alternatives with the coefficients $\boldsymbol{\beta_J}$ of the same variables estimated on the exhaustive set of alternatives. Two hypotheses are tested:

$$\begin{cases} H_0 : \beta_J = \beta_i \\ H_1 : \beta_J \neq \beta_i \end{cases}$$

Under $\boldsymbol{H_0}$, there is no systematic difference between the coefficients, indicating no correlation between the error term of the model and the regressors. In that case, random effects models are efficient. The alternative hypothesis corresponds to coefficients that differ. The hypothesis allows for correlation between the error term and the regressors, corresponding to fixed effects models, such as conditional logits.

The results of the Hausman test give a $\chi_2 = 13.88$ for 9 degrees of freedom and a p-value of 0.127 > 0.05. The null hypothesis cannot be rejected; thus, the IIA is not valid, indicating that random effects models are more efficient for our analysis. We therefore relax the IIA hypothesis and run a random parameter logit (or mixed logit) model.

3.1.2 Mixed logit results

As described in the section 2, the output of a mixed logit model is the estimated parameters of a distribution. In our study, we estimate the mean and standard deviation of the normal distribution of random parameters. Therefore, the absolute values of the estimates cannot be interpreted: a mean that is close to zero does not mean that the associated variable does not influence the choice but rather indicates heterogeneity in preferences. Some respondents might view the attribute as positive while others view it as negative.

Table 4 presents the results of the mixed logit model. All attributes, except for shape, significantly influence respondents' choices: individuals do not take into account whether a green space is linear. The size coefficient is slightly less significant (at 5%) than the coefficients associated with the other variables. The ASC coefficient estimate is high, indicating that other variables influence respondents' choices.

Signs of the estimated mean coefficients:

- tree density has a positive influence on the probability of choosing a given green space compared to grassland. However, the difference between going to a wooded area and a forest is not marked;
- the presence of water significantly increases respondents' utility;
- going to a large - greater than 1.5 ha - rather than a small green space positively

influences respondents' choices;

- compared to walking, all transportation modes, i.e., biking, driving and taking public transport negatively influence the choice of an urban green space;
- the utility associated with the status quo is negative.

Table 4 also shows that the standard deviations of most attributes (except for size) are significant, revealing heterogeneity in preferences. Individual coefficients vary significantly from the estimated mean.

Table 4: Mixed logit results

mean.woodland	1.256***	sd.woodland	0.441***
	(0.069)		(0.085)
mean.forest	1.231***	sd.forest	0.675***
	(0.075)		(0.084)
mean.shape	−0.073	sd.shape	0.113
	(0.055)		(0.102)
mean.water	0.675***	sd.water	0.496***
	(0.059)		(0.070)
mean.size	0.085*	sd.size	0.007
	(0.037)		(0.029)
mean.bike	−0.548***	sd.bike	0.559***
	(0.071)		(0.087)
mean.PTcar	−0.889***	sd.PTcar	0.930***
	(0.083		(0.082)
mean.time	−0.060***	sd.time	0.009**
	(0.003)		(0.003)
mean.ASC	−3.851***	sd.ASC	1.903***
	(0.204)		(0.161)
Observations	3,672		
Log Likelihood	−2,676.942		

Note: *p<0.1; **p<0.05; ***p<0.01

3.1.3. Willingness to accept longer travel time to reach a green space

With the mixed logit model, we account for individual heterogeneity with an individual random component in the utility expression. We assume that this random component is normally distributed among individuals. Because coefficients are individually estimated, we can represent the distribution of individual willingness to accept (WTA) to better express the heterogeneity of preferences.

Here, our "monetary" attribute is the time spent to reach a site. WTA is thus the willingness of respondents to accept a longer trip duration to reach a green space given the different attribute levels of the site.

From equation 3, we calculate the mean individual WTA for all significant attributes (Table 5). On average, people are willing to spend 21 min to travel to a woodland or a forest, compared to a grassland. This time is reduced to 11 min for the presence of a body of

water and 1 min for a green space larger than 1.5 ha. If the average respondent had to take a bike instead of walking, he would be willing to reduce the journey by 9 minutes and by 15 minutes if he had to take a car or public transport.

Table 5: WTA estimates

	Woodland	Forest	Water	Size	Bike	Public transport or Car
Average WTA	21 min	21 min	11 min	1 min	- 9 min	- 15 min

Figures 5a to 5g plot the kernel densities of the distribution of the individual parameter estimates (in black) and the corresponding normal distribution (in blue). The plotted distributions show that preferences do not follow a normal distribution, reflecting heterogeneity in preferences among individuals, except for the size attribute. This result is in line with the non-significant estimate of the standard deviation of the size attribute. For all other attributes, distinctive groups are observed, as is particularly evident for the ASC, for which two clear groups emerge. Except for the ASC, the plotted distributions reveal heterogeneity among individuals without a clear indication of how many groups exist. However, the majority of respondents have a negative preference for the status quo.

On the basis of the observation that different groups appear to emerge with respect to individual preferences, we attempted to specify these groups via a latent class model.

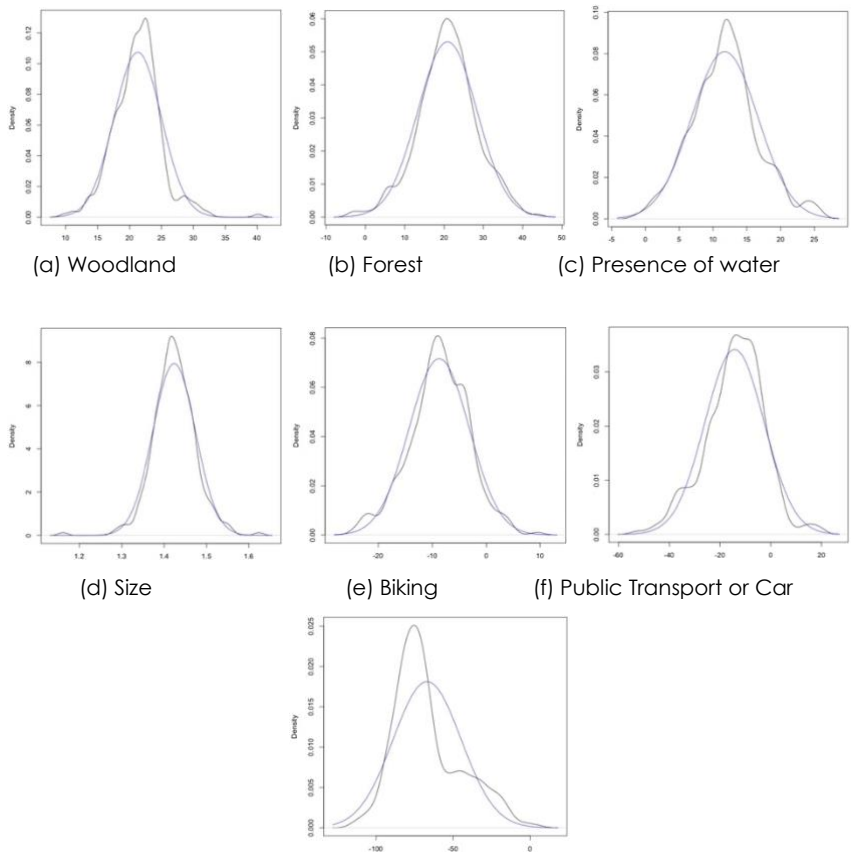


Figure 5: Individuals' willingness to accept distribution for each significant attribute (from subfigure a to f) and the ASC (g)

3.1.4 Latent class logit

To determine how many classes best fit our data, we obtained the Bayesian information criterion (BIC) and Akaike information criterion (AIC) for a latent class model with 2 and 3 classes. Given our limited sample size, considering more than 3 classes would result in an insufficient number of individuals in each class. According to the BIC, a model with 3 classes is more efficient than a 2-class model. We also determine the predictive quality of the model, thereby confirming the choice of a three-class model.

Our approach contains two steps. First, we determine preferences associated with each class by means of a standard latent class model. Then, we regress the posterior membership probability (as defined in equation A.5) on individual socio-economic variables. Our objective is to relate preferences to different individual-specific characteristics. Because of missing data concerning follow- up questions, our sample is reduced to 306 respondents, equivalent to 11 016 observations.

Figure 7 shows the latent class results, and 6 indicates the WTA for all attributes and for the three classes. Orthogonal preferences stand out in 2 ways: between forest cover and the presence of water and between transportation modes. The willingness to spend time travelling to a recreational green space varies by a factor of two between the first and third classes.

First-class individuals prefer woodlands that are 23 minutes from their home and prefer to walk. Such individuals will not choose a green space if they have to take public transport, a car or a bike. Among the three classes, first-class individuals display the lowest willingness to spend time travelling to a green space (23 minutes vs 42 minutes and 39 minutes for the second and third classes). Compared to the average commuting time in the region (68 min to travel to and from work), this value is high. This class is also looking for water, but to a lesser extent, and is willing to travel an additional 6 min for this attribute.

In contrast to the first and third classes, the second class does not value tree cover. Second-class individuals are heavily interested in water, preferably not in a linear form. These individuals are willing to travel 37 min to reach a green space with a body of water. They will neither take a car nor public transport but would take a bike (even though it would not positively influence their preference). One of their most preferred green spaces would therefore be ponds.

The third class definitely prefers woodlands and forests. People in this class are willing to spend 36 min to travel to a forest and 39 min to travel to wood- lands. Unlike the first two classes, these individuals do not mind taking a car or public transport. Their preference for water is limited, with a WTA of 5 min to travel to a green space with a body of water.

Overall, the first and third classes are definitely interested in dense tree cover compared to grassland. As the first class is only willing to travel by foot, individuals in this class can travel up to twenty minutes for a woodland or a forest, while those in the third class can double this time by taking cars or public transport. The second class has completely different preferences: the most important attribute is water and not tree cover. Second-class individuals would travel up to 42 min to find a green area with a body of water.

Table 6: Willingness to accept for each class

	Class 1	Class 2	Class 3
Woodland	23 min	*Not significant*	39 min
Forest	18 min	*Not significant*	36 min
Shape *(Linear)*	*Not significant*	- 13 min	*Not significant*
Water	6 min	42 min	5 min
Size	*Not significant*	7 min	*Not significant*
Bike	- 24 min	*Not significant*	- 7 min
Public Transport or Car	- 18 min	- 47 min	*Not significant*

3.1.5. Posterior class-membership regression

We chose to regress posterior class membership on covariates for which we had sufficient data: socio-professional category, age, revenue, diploma, main declared activity when visiting a green space, department of residence, number of visits to a green space last year and having a private garden. Table 8 presents the reference categories.

Table 7: Latent class results

	Class 1	Class 2	Class 3
Woodland	1.065***	0.022	1.691***
	(0.163)	(0.134)	(0.220)
Forest	0.824***	-0.069	1.574***
	(0.168)	(0.170)	(0.174)
Shape	0.034	-0.203**	-0.026
	(0.111)	(0.077)	(0.091)
Water	0.262*	0.717***	0.209*
	(0.117)	(0.096)	(0.090)
Size	0.186	0.163*	0.063
	(0.117)	(0.080)	(0.089)
Bike	-1.128***	-0.171	-0.283 ·
	(0.176)	(0.122)	(0.164)
Public transport or Car	-0.834***	-0.858***	-0.213
	(0.146)	(0.122)	(0.141)
Transportation time	-0.046***	-0.019***	-0.043***
	(0.006)	(0.005)	(0.005)
ASC	-0.798***	-3.595***	-3.798***
	(0.225)	(0.363)	(0.549)

Significance: $*** \equiv p < 0.001$; $** \equiv p < 0.01$; $* \equiv p < 0.05$; $\cdot \equiv p < 0.1$

Table 8: Reference categories for the logit regression

Variable	Reference category
Socio-professional category	Other without professional activity
Revenue	$< 1400€$
Age	[45,59]
Diploma	French Certificate of General Education

Table 9 synthesizes all dummy-coded variables that we tested for influence on class membership.

Appendix G shows the odds ratios (as defined in equation D.3). Concerning the influence of socio-professional categories, which are highly correlated with revenues, the results give counter-intuitive information. For class 1, only one socio-professional category is significant, whereas all revenue groups (compared to the reference category) are significant. In class 2, significant socio- professional groups include employee to higher-educated individuals and inter- mediate professionals, whereas no revenue category is significant. One explanation is that revenue and socio-professional categories are sensitive information, which means that respondents worry about the repercussions associated with their answer (Barnett, 1998), resulting in biased information. Apart from this enigmatic result, 3 different membership profiles stand out:

- Class 1: living in Paris, or in the little crown, being quite wealthy (earning more than 2000€ monthly) and aged more than 45 increase class-1 membership probability.
- Class 2: being young (18 to 44 years old) and doing activities in contact with nature (such as fishing, observation, walking, refreshing) positively influence class-2 membership probability.
- Class 3: having kids, pets, being a woman and looking for open-air activities positively influence class-3 membership probability.

Class-2 and Class-3 contain people from the big crown. Overall, department of residence, age and green space usage are the main variables that influence preferences.

4. Discussion

By relating preference heterogeneity to socio-economic variables, our application of a distance-based DCE provides information to adjust urban planning

Table 9: Description of the logit regression variables (dummy-coded)

	Description
Commercant	Being a retailer, a craftsman or an entrepreneur
Cadre	Having a managerial or a high-education profession
Prof_inter	Having an intermediate profession
Employe	Being an employee
Ouvrier	Being a worker
Retraite	Being a retiree
Chien	Main activity when going to a green space: walking the dog
Sport	Main activity when going to a green space: doing sports
Frais	Main activity when going to a green space: cooling off
Kids	Main activity when going to a green space: enjoying equipment for kids
Nature	Main activity when going to a green space: observing nature
Peche	Main activity when going to a green space: fishing
Pn	Main activity when going to a green space: picnic
Relax	Main activity when going to a green space: relaxing
Marche	Main activity when going to a green space: walking
Rev2	Having a disposable monthly revenue between 1400€ and 2000€
Rev3	Having a disposable monthly revenue between 2000€ and 3000€
Rev4	Having a disposable monthly revenue exceeding 3000€
CAP	Last diploma is: CAP (Certificate of professional competence)
Bac	Last diploma is: French baccalaureate
Sup	Last diploma is of higher-education
Petite_couronne	Living in the "small crown": departments 92, 93, 94
Paris	Living in Paris "intramuros"
Jeunes	Aged [18,29]
Jeunes_adultes	Aged [30,44]
Adultes_plus	Aged [60,74]
Seniors	Aged older than 75
nb_vis_lastyear	Last year's number of visits to a green space
private_garden	Having a private garden or not

policies. Green space recreational preferences are influenced by place of residence, income, age, and activities practiced when visiting a green space.

Parisians, and to a lesser extent inhabitants from the little crown, especially when older than 45 years, prefer woodlands and forests within a 23-min trip. As these individuals reject any other transportation mode than walking, this trip corresponds to 1380 meters (for a person walking at 3.6 km per hour). The size of the green space does not matter, but these individuals reject biking and taking a car or public transport. Therefore, a complement to the conversion of artificialized land to green spaces would be to increase tree cover in existing parks (La Vilette for example) or in the streets. Figure 2 shows that in 11 districts, less than 50% of the population has access to a green area of at least 1.5 ha less than 300 meters away. In those districts, such policies would increase population well-being. However, the prioritization among municipalities for renaturation actions should include other components of well-being (e.g., material living conditions, health, education,

personal activities, political participation, economic and physical insecurity) to ensure fair decision making (Liotta and al., submitted).

Big crown young residents, aged between 18 and 44 years and who like being in contact with nature (i.e., observing nature, walking or fishing) choose their recreational green space based on the presence of a body of water. These individuals are willing to spend 42 min to travel to such a green space. An efficient policy to target this audience in these departments would be to restore ponds or to reopen waterways (such as La Bièvre, a portion of which reopened in 2016), and develop biking lanes (which could also support sustainable mobility).

Modest families (with a disposable income between 1400€and 2000€monthly), especially those with pets, are willing to travel a 39 min for a woodland. They consider taking a car or public transport, which corresponds to a 26 km travel distance (at a 50 km per hour), but not biking. The size of the green space does not alter their willingness to spend time reaching a site. An efficient policy to target this portion of the population, without any land-use change, would be to increase accessibility to green spaces to ensure a maximum 39-min drive/public transportation, for example, by increasing entries and exits. Regardless of the green space sizes, increasing tree cover in artificialized open spaces and adjusting green space facilities to the activities families engage in (allow dogs if they are not allowed, develop kid playground or enhance possible fishing areas) would increase individuals' well-being.

Our study is subject to the limits of stated preferences methods, including hypothesis bias (when respondents have difficulty putting themselves in fictive situations) and anchoring bias (respondents rely too much on the first information they obtain when they make a decision). Therefore, our results should be put into perspective of the interpretation respondents made of the questionnaire and the quality of their responses. Moreover, even if the sample is representative of the general population of Ile-de-France because of the sampling strategy and the weights applied in the econometric models, the sample remains small. An increase in the size of the survey sample may slightly change the conclusions.

Our results provide information to decision makers to satisfy the actual preferences regarding urban nature in Ile-de-France. Although these results are useful for calibrating land-use planning in urban areas, renaturation and conservation policies cannot be measured by the sole channel of recreational service. One ongoing study is exploring other ecosystem services, such as urban heat island mitigation, natural heritage and water retention (Tardieu et al., in progress). Finally, this study ignores the spatial connection between green spaces, which may have a strong influence on individuals' preferences to practice some activities such as running or cycling. For instance, STRAVA data shows that francilians cover different types of green spaces when they run[12]. This aspect should be integrated into further research.

This study provides additional material for the land-sparing/land-sharing debate. This debate was initiated by the agricultural economics literature but can be applied in our case. Indeed, in the context of rising food demand and in- creased pressure on biodiversity, land-sparing strategies consist of keeping high- yield farming separated from natural habitats that are protected from conversion to agriculture. Land-sharing strategies consist of having both agriculture production and conservation goals on the same land (Phalan et al., 2011).

As individuals in the third class are the only ones to consider driving/taking public transport, for a significant travel time, maintaining separate land dedicated to conservation would

[12] More information on where Parisians run: https://www.institutparisregion.fr/fileadmin/DataStorage/SavoirFaire/vignettescartes/Societe/CE_25_Parcours_ sportifs.jpg

be in line with individuals' preferences in terms of recreational services. Land-sparing policies could therefore be implemented. By contrast, in the first two classes, individuals are less willing to travel long distances. Conservation/renaturation objectives should therefore be integrated on the same lands as food production, and areas advocating for land-sharing policies should be implemented. In Ile-de-France, resident preferences indicate that urban policies must balance land-sharing in Paris intra-muros and the little crown and land-sparing in the big crown.

Appendix A. Method specification: conditional logit

Following the approaches of Lancaster and McFadden, if y is a binary variable indicating whether an alternative is chosen (y = 1) or not (y = 0), we have for individual n choosing alternative i over any other alternative *j*:

$$\begin{cases} y_{n,i} = 0 \Leftrightarrow \forall j \neq i,\ U_{n,j} > U_{n,i} \\ y_{n,i} = 1 \Leftrightarrow \forall j \neq i,\ U_{n,i} > U_{n,j} \end{cases}$$

Replacing with the expression of equation 1:

$$\begin{aligned} \mathbb{P}(y_{n,i} = 1 \mid X_i) &= \mathbb{P}\big(V_{n,i} + \epsilon_{n,i} > V_{n,j} + \epsilon_{n,j}\big) \\ &= \mathbb{P}\big(\epsilon_{n,j} < \epsilon_{n,i} + V_{n,i} - V_{n,j}\big) \end{aligned}$$

Individual n will chose alternative i among other alternatives *j* if the difference between the observed part of the utility plus the unobserved part of the utility from choosing alternative i is greater than the unobserved utility from any other alternative *j*.

Since the error terms follow a Gumbel distribution, $\varepsilon_{n,i}$ is characterized by its density and distribution functions, given by equations A.1 and A.2:

$$f(\epsilon_{n,i}) = e^{-\epsilon_{n,i}} e^{-e^{-\epsilon_{n,i}}} \qquad \text{(A.1)}$$

$$F(\epsilon_{n,i}) = \exp(-e^{-\epsilon_{n,i}}) \qquad \text{(A.2)}$$

Given that the $\varepsilon_{n,I}$'s are independent, we have:

$$\mathbb{P}(y_{n,i} = 1 \mid \epsilon_{n,i}) = \prod_{j \neq i} \exp(-e^{-(\epsilon_{n,i} + V_{n,i} - V_{n,j})}) \qquad \text{(A.3)}$$

Then, $P_{n,i}$, defined by $P_{n,i} = P(y_{n,i} = 1)$, is calculated as the integral of A.3 over the distribution of $\varepsilon_{n,i}$:

$$\mathbb{P}_{n,i} = \mathbb{P}(y_{n,i} = 1) = \int \left(\prod_{j \neq i} e^{-e^{-(V_{n,i} - V_{n,j} + \epsilon_{n,i})}} \right) e^{-\epsilon_{n,i}} e^{-e^{-\epsilon_{n,i}}} \partial \epsilon_{n,i} \qquad \text{(A.4)}$$

Let $X_{n,i}$ be the attribute levels of alternative i chosen by individual n. We postulate that utility is linear and can thus be written as:

$$V_{n,i} = X_{n,i} \beta_i \qquad \forall i$$

Equation A.4 can be written 13 as:

$$\mathbb{P}_{n,i} = \frac{\exp(\beta' X_{n,i})}{\sum_j \exp(\beta' X_{n,j})} \tag{A.5}$$

The conditional logit model is estimated by maximum likelihood. The log- likelihood function is given by:

$$LL(\beta) = \sum_{n=1}^{N} \sum_{i} y_{n,i} \ln P_{n,i}$$

Using equation A.5:

$$LL(\beta) = \sum_{n=1}^{N} \sum_{i} y_{n,i} \ln \frac{\exp(\beta' X_{n,i})}{\sum_j \exp(\beta' X_{n,j})} \tag{A.6}$$

The model estimates the value of β that maximizes equation A.6.

Appendix B. Model specification: mixed logit

For a given value of ***βi***, we have:

$$\mathbb{P}_{n,i} \mid \beta_i = \frac{\exp(\beta_i' X_{n,i})}{\sum_j \exp(\beta_i' X_{n,j})} \qquad \text{(B.1)}$$

The unconditional probability is therefore:

$$\mathbb{P}_{n,i} = \int \frac{\exp(\beta' X_{n,i})}{\sum_j \exp(\beta' X_{n,j})} f(\beta \mid \theta) d\beta \qquad \text{(B.2)}$$

The integral B.2 is estimated by the maximum simulated likelihood estimator.

In each step of the optimization, ***R*** draws of the parameters ***β*** are taken from the assumed distribution, and the corresponding likelihood is computed as

$$L_{n,i}(\beta) = \frac{\exp(\beta' X_{n,i})}{\sum_j \exp(\beta' X_{n,j})}$$

The simulated probability corresponds to the average results of the R draws:

$$\hat{\mathbb{P}}_{n,i} = \frac{1}{R} \sum_R L_{n,i}(\beta_r)$$

with ***βr*** being the ***r-th β*** draw. The parameters θ of the underlying distribution (mean and standard deviation for a normal distribution) are optimized: the values of θ that maximize the simulated likelihood are the estimated parameters.

The calculation is detailed in Train (2009), p.74-7

Appendix C. Model specification: latent class

From equation A.5, we calculate the joint density of all choices made by individual n, when choosing alternative j among T choice situations, conditional on a specific value of $\beta i = \beta q$

$$f(Y_n \mid X_n, \beta_q) = \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\frac{\exp(X_{njt}^{\top} \beta_q)}{\sum_{j=1}^{J} \exp(X_{njt}^{\top} \beta_q)} \right]^{Y_{njt}} \qquad \text{(C.1)}$$

where $Yn = (yn1, yn2, ..., ynT)$ is a vector of the sequence of choices made by 600 individual n.

For a given number of classes **Q**, the membership probability of each class $W_{nq}\, i$

$$w_{nq}(\gamma) = \frac{\exp\left(h_n^{\top}(\gamma_q)\right)}{\sum_{q=1}^{Q} \exp\left(h_n^{\top}(\gamma_q)\right)};$$

where h_n is a set of socio-economic variables that influence class membership.

A priori, we hypothesize that no socio-economic variable determines class membership. We thus have a constant class membership among individuals with $q=1...Q$; $\sum_q w_{nq} = 1$; $w_{nq} > 0$ and $\gamma_1 = 0$

From equation C.1, we obtain the unconditional probability of the sequence of choices made by individual n as:

$$f(Y_n \mid X_n, \theta) = \sum_{q=1}^{Q} w_{nq}(\gamma_q) \left\{ \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\frac{\exp(x_{njt}^{\intercal} \beta_q)}{\sum_{j=1}^{J} \exp(x_{njt}^{\intercal} \beta_q)} \right]^{Y_{njt}} \right\} \qquad \text{(C.3)}$$

Where $\theta = (\gamma, \beta)$ is the vector of parameters at the population level such that $\beta = (\beta 1, ..., \beta Q)$ and $\gamma = (\gamma 1, ..., \gamma Q)$.

Individual-specific estimates. The discrete distribution of the parameters gives an overview of how preferences are distributed in the population. To determine where an individual is situated in the distribution given a set of choices, we calculate individual-specific parameters. We estimate the posterior membership probability as follows (Greene and Hensher, 2003; Kamakura and Russell, 1989):

$$\hat{\pi}_{nq}(\beta_n \mid Y_n, X_n, \theta) = \frac{\hat{w}_{nq}(\gamma_q) \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\frac{\exp(X_{njt}^{\intercal} \hat{\beta}_q)}{\sum_{j=1}^{J} \exp(X_{njt}^{\intercal} \hat{\beta}_q)} \right]^{Y_{njt}}}{\sum_{q=1}^{Q} \hat{w}_{nq}(\gamma_q) \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\frac{\exp(X_{njt}^{\intercal} \hat{\beta}_q)}{\sum_{j=1}^{J} \exp(X_{njt}^{\intercal} \hat{\beta}_q)} \right]^{Y_{njt}}} \qquad \text{(C.4)}$$

which is the probability that individual n belongs to class q given the observed choices

Appendix D. Relating class membership to auxiliary variables

We regress individuals' logit-transformed posterior probability (as defined in equation A.5) on auxiliary variables. Formally, the logit regression specification is:

$$\log\left(\frac{\hat{\pi}_{nq}}{1-\hat{\pi}_{nq}}\right) = \beta_1 X_1 + ... + \beta_K X_K \qquad \text{(D.1)}$$

where ***Xk= (x1,...,xK)*** is the set of auxiliary variables. From equation A.6, we have the odds expression

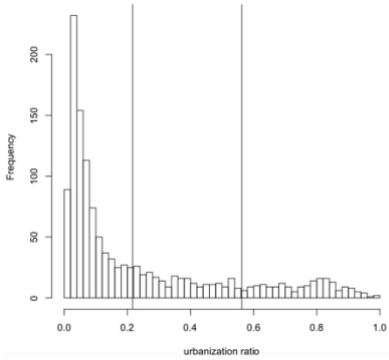
$$\text{Odds} = \frac{\hat{\pi}_{nq}}{1-\hat{\pi}_{nq}} = \exp(\beta_1 X_1 + ... + \beta_K X_K) \qquad \text{(D.2)}$$

To evaluate the influence of a variable ***Xk***, we calculate the odds ratio when the variable increases by 1 unit:

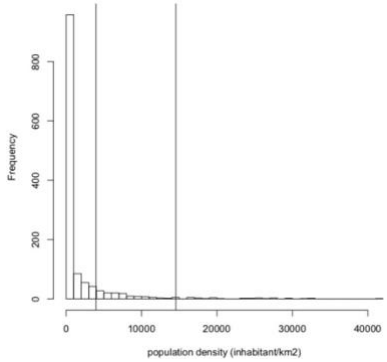
$$\text{Odds ratio} = \frac{Odds_{X_k+1}}{Odds_{X_k}} = \exp(\beta_k) \qquad \text{(D.3)}$$

Thus, if exp***(βk) > 1, Xk*** has a positive influence on the class Q membership probability.

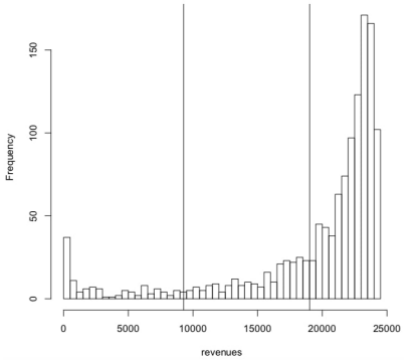
Appendix E. Jens class results for clustering municipalities



(a) Clusters of municipalities according to



(b) Clusters of municipalities according to their urbanization ratio their population density



(c) Clusters of municipalities according to their revenue

Appendix F. Questionnaire example

T1

Questionnaire on green spaces in Ile-de-France

Hello,

We are conducting a survey in Ile-de-France to better understand people's preferences concerning green spaces (parks, forests, riverbanks, etc.) in Ile-de-France.

Would you please have a few minutes to answer the questionnaire? *(10 minutes maximum)*

The questionnaire is **anonymous**.

Only persons over 18 years-old are allowed to answer

Profile

1) What is your place of residence?

CITY	POSTAL CODE
_____	_____

2) Are you a man or a woman?

☐ Man ☐ Woman

3) How old are you?

☐ 18 – 29 years-old
☐ 30 – 44 years-old
☐ 45 – 59 years-old
☐ 60 – 74 years-old
☐ 75 years-old and older

4) What is your socio-professional category?

☐ Farmer
☐ Craftsman, Retailer, Entrepreneur
☐ Managerial and higher-education profession
☐ Intermediate profession
☐ Employee
☐ Worker
☐ Retiree
☐ Other without professional activity

5) How many adults and children (without age limit) constitute your household ?

ADULT(S)	CHILD/CHILDREN (if any)
_____	_____

Green space attributes	Alternative 1	Alternative 2	Status quo
Forest cover			I choose neither of both green spaces
Shape			
Presence of a water body			
Size	Large* / Long*	Small** / Short**	
Transport mode	Biking / Trotinette	By foot	
Transportation time	5 min	30 min	
Choice			1

Green space attributes	Alternative 1	Alternative 2	Status quo
Forest cover			I choose neither of both green spaces
Shape			
Presence of a water body			
Size	Small** / Short**	Large* / Long*	
Transport mode	Biking / Trotinette	Public Transport/Car	
Transportation time	5 min	30 min	
Choice			2

***Large/Long:** corresponds to a green space larger than 1.5ha (100x150m) or longer than 1.5 km

****Small/Short**: corresponds to a green space smaller than 1.5ha (100x150m) or shorter than 1.5 km

Green space attributes	Alternative 1	Alternative 2	Status quo
Forest cover			I choose neither of both green spaces
Shape			
Presence of a water body			
Size	Large* / Long*	Small** / Short**	
Transport mode	By foot	Biking / Trotinette	
Transportation time	10 min	20 min	
Choice			3

Green space attributes	Alternative 1	Alternative 2	Status quo
Forest cover			I choose neither of both green spaces
Shape			
Presence of a water body			
Size	Small** / Short**	Large* / Long*	
Transport mode	By foot	Public Transport/Car	
Transportation time	20 min	10 min	
Choice			4

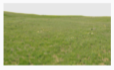
Green space attributes	Alternative 1	Alternative 2	Status quo
Forest cover			I choose neither of both green spaces
Shape			
Presence of a water body			
Size	Small** / Short**	Large* / Long*	
Transport mode	Public Transport/Car	Biking / Trotinette	
Transportation time	30 min	5 min	
Choice			5

Green space attributes	Alternative 1	Alternative 2	Status quo
Forest cover			I choose neither of both green spaces
Shape			
Presence of a water body			
Size	Small** / Short**	Large* / Long*	
Transport mode	By foot	Biking / Trotinette	
Transportation time	10 min	20 min	
Choice			6

Green space attributes	Alternative 1	Alternative 2	Status quo
Forest cover			I choose neither of both green spaces
Shape			
Presence of a water body			
Size	Small** / Short**	Large* / Long*	
Transport mode	Public Transport/Car	By foot	
Transportation time	10 min	20 min	
Choice			7

Green space attributes	Alternative 1	Alternative 2	Status quo
Forest cover			I choose neither of both green spaces
Shape			
Presence of a water body			
Size	Large* / Long*	Small** / Short**	
Transport mode	Public Transport/Car	Biking / Trotinette	
Transportation time	5 min	30 min	
Choice			8

Green space attributes	Alternative 1	Alternative 2	Status quo
Forest cover			I choose neither of both green spaces
Shape			
Presence of a water body			
Size	Large* / Long*	Small** / Short**	
Transport mode	Biking / Trotinette	Public Transport/Car	
Transportation time	30 min	5 min	
Choice			9

Green space attributes	Alternative 1	Alternative 2	Status quo
Forest cover			I choose neither of both green spaces
Shape			
Presence of a water body			
Size	Large* / Long*	Small** / Short**	
Transport mode	By foot	Public Transport/Car	
Transportation time	20 min	10 min	
Choice			10

Green space attributes	Alternative 1	Alternative 2	Status quo
Forest cover			I choose neither of both green spaces
Shape			
Presence of a water body			
Size	Small** / Short**	Large* / Long*	
Transport mode	Biking / Trotinette	By foot	
Transportation time	20 min	10 min	
Choice			11

Green space attributes	Alternative 1	Alternative 2	Status quo
Forest cover			I choose neither of both green spaces
Shape			
Presence of a water body			
Size	Large* / Long*	Small** / Short**	
Transport mode	Public Transport/Car	By foot	
Transportation time	30 min	5 min	
Choice			12

Have you systematically omitted one or several attribute(s) in the choices you had to make? If yes, which one?

Respondent's type

1) **How many times have you been to a green space (park, forest, riverbank, garden...) last year (March 2018 to March 2019)?**

2) **What is the name of the green space you visited the most last year (March 2018 to March 2019)?**

3) **What activities do you usually do when going to a green space? (Rank from 1 to 3)**

- ☐ Walk the dog
- ☐ Practice sports
- ☐ Cool-off
- ☐ Enjoy kids' facilities
- ☐ Observe nature
- ☐ Fishing / Hunting
- ☐ Have a picnic
- ☐ Relax
- ☐ Walk
- ☐ Other: .....

4) **How long do you stay in average in a green space?**

5) **How long it took you to go to a green space, in average, last year (March 2018 to March 2019)?**

6) **What facilities do you prefer when going to a green space?**

- ☐ None
- ☐ Benches

☐ Lights
☐ Sports facilities
☐ Playground for kids
☐ Food and beverage facilities
☐ Multiple entries and exists
☐ Other: ……

7) In average last year (March 2018 to March 2019), how much did you spend for your green spaces' excursion (sport equipment, binoculars, fishing rod…)?

8) Did you feel secure the last time you went to a green space?

☐ Yes
☐ No

9) What are the 3 main activities that you do when leaving your house?

1: 2: 3:

10) Do you have a private garden?

☐ Yes
☐ No

11) Do you feel concern by environmental issues?

☐ Yes, a lot
☐ Yes but not more than other issues
☐ Not that much
☐ Not at all

6) Which income range do you fall under (net, but including social services)?

☐ Strictly lower than 1 400 euros
☐ Between 1 400 and 2 000 euros
☐ Between 2 000 and 3 000 euros
☐ More than 3 000 euros

7) What is the highest diploma you obtained?

☐ None, French Certificate of general Education
☐ Certificate of professional competence or Higher technical certificate
☐ Baccalaureate (general, technological or professional)
☐ Higher education diploma

Interview

Place of the interview:

Date and hour :

Appendix G. Odds ratios for the 3 identified classes

Appendix G. Odds ratios for the 3 identified classes

	Class 1	Class 2	Class 3
Commercant	**5.71**	0.12	0.24
Cadre	0.24	**4.91**	0.53
Prof_inter	0.54	**6.88**	0.13
Employe	0.57	**3.21**	0.39
Ouvrier	*Not significant*	*Not significant*	*Not significant*
Retraite	0.24	0.49	**2.53**
Chien	0.01	*Not significant*	**3.45**
Sport	0.02	0.41	*Not significant*
Frais	0.006	**1.65**	*Not significant*
Kids	0.003	**1.45**	**7.35**
Nature	0.007	**5.73**	*Not significant*
Peche	2.98e-05	**3.44**	*110.46*
Pn	0.03	**1.71**	0.27
Relax	0.02	0.62	**1.74**
Marche	0.01	**2.39**	*Not significant*
Rev2	**2.72**	0.37	**1.46**
Rev3	**6.58**	0.14	*Not significant*
Rev4	**4.69**	*Not significant*	0.38
CAP	0.29	0.51	**2.86**
Bac	**4.86**	0.11	0.27
Sup	**1.83**	0.28	*Not significant*
Petite_couronne	**2.31**	0.21	*Not significant*
Paris	**7.14**	0.36	0.20
Jeunes	*Not significant*	**1.64**	*Not significant*
Jeunes_adultes	0.66	**1.74**	*Not significant*
Adultes_plus	**5.10**	*Not significant*	0.04
Seniors	**1.48**	0.33	0.02
nb_vis_lastyear	*Not significant*	0.99	0.99
private_garden	**1.96**	*Not significant*	0.13
gender	0.15	0.64	**2.09**

Acknowledgements

This research was funded by ADEME, MTES and AgroParistech. This study is part of the IDEFESE project (https://idefese.wordpress.com/). We thank Claire Benedetti, Claire Lefebvre, Victor Leon, Lucie Thiolliere, Baptiste Wozniak for their precious help in the interviews. We also thank all the stakeholders involved in the IDEFESE project for their useful feedbacks, and in particular the Ile-de-France region and the Institut Paris Region.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Adamowicz, W., & Boxall, P. (2001). Future directions of stated choice methods for environment valuation. Paper prepared for 'Choice Experiments: A New Approach to Environmental Valuation Conference'.

Adamowicz, W., Louviere, J., & Williams, M. (1994). Combining revealed and stated preference methods for valuing environmental amenities. Journal of environmental economics and management, 26, 271–292.

Alcock, I., White, M. P., Wheeler, B. W., Fleming, L. E., & Depledge, M. H. (2014). Longitudinal effects on mental health of moving to greener and less green urban areas. Environmental science & technology, 48 , 1247–1255.

Barnett, J. (1998). Sensitive questions and response effects: an evaluation. Journal of Managerial Psychology 13, 63–76.

Cabral, P., Feger, C., Levrel, H., Chambolle, M., & Basque, D. (2016). Assessing the impact of land-cover changes on ecosystem services: A first step toward integrative planning in Bordeaux, France. Ecosystem Services, 22, 318–327.

Chakir, R., David, M., Gozlan, E., & Sangare, A. (2016). Valuing the impacts of an invasive biological control agent: a choice experiment on the Asian ladybird in France. Journal of agricultural economics 67, 619–638.

Clark, S. L., & Muthén, B. (2009). Relating latent class analysis results to variables not included in the analysis.

Coppel, G., & Wüstemann, H. (2017). The impact of urban green space on health in berlin, Germany: Empirical findings and implications for urban planning. Landscape and Urban Planning, 167, 410–418.

Cox, D., Shanahan, D., Hudson, H., Fuller, R, Anderson, K., Hancock, S., & Gaston, K. (2017). Doses of nearby nature simultaneously associated with multiple health benefits. International journal of environmental research and public health, 14, 172.

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.

Espinosa-Goded, M., Barreiro-Hurlé, J., & Ruto, E. (2010). What do farmers want from agri-environmental scheme design? a choice experiment approach. Journal of Agricultural Economics, 61, 259–273.

Green, P. E., & Rao, V. R. (1971). Conjoint measurement-for quantifying judgmental data. Journal of Marketing research, 8, 355–363.

Green, P. E., Wind, Y., & Jain, A. K. (1972). Consumer menu preference: An application of additive conjoint measurement. ACR Special Volumes.

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research Part B: Methodological, 37, 681–698.

Hanley, N., Wright, R. E., & Adamowicz, W. (1998). Using choice experiments to value the environment. Environmental and resource economics, 11, 413– 428.

Kamakura, W. A., & Russell, G. J. (1989). A probabilistic choice model for market segmentation and elasticity structure. Journal of marketing research, 26, 379–390.

Kermagoret, C., Levrel, H., Carlier, A., & Dachary-Bernard, J. (2016). Individual preferences regarding environmental offset and welfare compensation: a choice experiment application to an offshore wind farm project. Ecological Economics, 129, 230–240.

Kontoleon, A., & Yabe, M. (2003). Assessing the impacts of alternative 'optout' formats in

choice experiment studies: consumer preferences for genetically modified content and production information in food. Journal of Agricultural policy and Resources, 5, 1–43.

Lancaster, K. J. (1966). A new approach to consumer theory. Journal of political economy, 74, 132–157.

Levrel, H., Cabral, P., Feger, C., Chambolle, M., & Basque, D. (2017). How to overcome the implementation gap in ecosystem services? a user-friendly and inclusive tool for improved urban management. Land Use Policy, 68, 574–584.

Liotta, C., Kervinio, Y., Levrel, H., & Tardieu, L. (). Targeting renaturation: a constructive approach for fair decision making. Submitted to Environmental Science and Policy.

Luck, G. W., Davidson, P., Boxall, D., & Smallbone, L. (2011). Relations between urban bird and plant communities and human well-being and connection to nature. Conservation Biology, 25, 816–826.

Markevych, I., Tiesler, C. M., Fuertes, E., Romanos, M., Dadvand, P., Nieuwenhuijsen, M. J., Berdel, D., Koletzko, S., & Heinrich, J. (2014). Access to urban green spaces and behavioural problems in children: Results from the giniplus and lisaplus studies. Environment international, 71, 29–35.

McFadden, D. et al. (1973). Conditional logit analysis of qualitative choice behavior,

Mulatu, D. W., Alvsilver, J., & Siikamäki, J. (2019). Valuing residents' preferences for improved urban green space ecosystem services in Addis Ababa, Ethiopia.

Niemelä, J., Saarela, S.-R., Sôderman, T., Kopperoinen, L., Yli-Pelkonen, V., Väre, S., & Kotze, D. J. (2010). Using the ecosystem services approach for better planning and conservation of urban green spaces: a Finland case study, Biodiversity and Conservation, 19, 3225–3243.

Phalan, B., Onial, M., Balmford, A., & Green, R. (2011). Reconciling food production and biodiversity conservation: Land sharing and land sparing compared. Science (New York, N.Y.), 333, 1289–91. doi:10.1126/science.1208742.

Tardieu, L., Coste, L., Hamel, P., Viguie, V., & Levrel, H. (). Does the ecosystem services approach provide additional information to urban planners compared to simple land use change analysis? Forthcoming.

Train, K. E. (2009). Discrete choice methods with simulation. Cambridge university press.

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.

Vaissière, A.-C., Tardieu, L., Quétier, F., & Roussel, S. (2018). Preferences for biodiversity offset contracts on arable land: a choice experiment study with farmers. European Review of Agricultural Economics, 45, 553–582.

Vermunt, J. (2010). Latent class modeling with covariates: Two improved three- step approaches, 18. doi:10.2307/25792024.