# Eliciting Consequentiality in Stated Preference Surveys: An Application to Urban Green

Ewa Zawojska*
*University of Warsaw, Faculty of Economic Sciences, Warsaw, Poland*

Malte Welling
*Institute for Ecological Economy Research, Berlin, Germany*

Julian Sagebiel
*Technische Universität Berlin, Institute for Landscape Architecture and Environmental Planning, Berlin, Germany*

* Corresponding author: ezawojska@wne.uw.edu.pl

**Abstract**: Information on respondents' perceptions about survey consequentiality is typically collected close to the end of the survey, following the preference elicitation. We inquire whether—and if so, how—the location and a repetition of a consequentiality perception elicitation question matter for stated consequentiality perceptions and for stated preferences. To that end, we use data from a discrete choice experiment survey conducted in Germany, in which respondents evaluated a project of expanding urban green areas. The survey involved two treatments: in one, respondents were asked about their consequentiality perceptions only after completing all preference elicitation tasks; in the other, respondents' consequentiality perceptions were elicited twice, that is, before and after the preference elicitation tasks. Based on ordered logit models, we find that stated consequentiality is stronger (i) when respondents are asked about consequentiality perceptions twice (that is, before and after preference elicitation tasks) rather than when asked only once (that is, after the tasks); and (ii) when the perceptions are inquired before the preference elicitation tasks rather than when inquired after the tasks. The latter finding emerges in both within-sample and across-sample comparisons. Mixed logit willingness-to-pay-space models reveal that stated willingness-to-pay values for project attributes significantly differ depending on whether the consequentiality elicitation question is asked before the preference elicitation or not. Furthermore, the consequentiality perceptions appear to significantly affect stated willingness-to-pay values, and this influence differs depending on whether the consequentiality elicitation question is asked before or after preference elicitation tasks. The findings of differences in the impact of the way consequentiality perceptions are elicited on willingness-to-pay values of attributes are observed for each of four German cities in which the survey was conducted. This evidence suggests that when willingness-to-pay values are corrected based on stated consequentiality perceptions—as often done in recent stated preference studies—the values may be sensitive to the way the consequentiality perceptions are elicited.

**Keywords**: stated preferences, discrete choice experiment, consequentiality perceptions, urban green spaces, Germany

# 1 Introduction

Stated preference surveys are used to learn about the value of public goods to the society. The methodology is applied in many areas, including environmental economics. The value estimates are needed, among others, for the assessment of benefits from the public good provision in cost-benefit analyses and for the evaluation of losses in litigation processes over natural damages. In order to incentivise respondents to reveal their preferences truthfully in such surveys, it is suggested that the surveys should be (viewed by respondents as) consequential; that is, the survey results should be seen as potentially influencing actual decisions regarding the public good provision and the related payment collection (Carson and Groves, 2007; Johnston et al., 2017). Despite consequentiality being acknowledged as a necessary condition for truthful preference disclosure (e.g., Carson, Groves and List, 2014; Vossler, Doyon and Rondeau, 2012), research and guidance on how to elicit respondents' perceptions about survey consequentiality are scarce (Lloyd-Smith, Adamowicz and Dupont, forthcoming). Information on the perceptions is typically collected via a follow-up question(s) after preference elicitation (e.g., Czajkowski et al., 2017; Oehlmann and Meyerhoff, 2017; Zawojska, Barczak and Czajkowski, 2019). To the best of our knowledge, no study to date has tested whether the location of the consequentiality perception elicitation question(s)—before or after preference elicitation—affects stated consequentiality perceptions and stated preferences in a discrete choice experiment survey involving a sequence of preference elicitation tasks.[1] This paper aims to address this research gap.

To comply with the consequentiality recommendation, two approaches are generally employed: researchers communicate potential real-life consequences of the survey result in the survey script (e.g., Andor, Frondel and Vance 2017; Drichoutis et al., 2017) and/or ask respondents about their perceptions regarding the survey consequentiality (e.g., Groothuis et al., 2017; Interis and Petrolia, 2014). In many field surveys, however, it might be difficult to credibly assure respondents about the survey consequentiality via the script so that their consequentiality perceptions are indeed affected (e.g., Czajkowski et al., 2017; Lloyd-Smith, Adamowicz and Dupont, forthcoming). Furthermore, in some cases, consequentiality needs to be deliberately kept vague, for example, when the presented project is at a hypothetical stage and policy-makers prefer not to make definite (consequential) statements in the survey. In the light of these difficulties, elicitation of consequentiality perceptions could sometimes appear as a more suitable approach, with a broader application, instead of inducing consequentiality via the survey script. Nevertheless, unanswered questions exist regarding the correct way of eliciting the perceptions.

A largely dominating method for eliciting consequentiality perceptions in stated preference surveys is to ask respondents a question of a type: "how likely do you think it is that the results of surveys such as this one will affect decisions about [the project presented]?" (Herriges et al., 2010, p. 74). A question such as this is usually answered on a Likert scale including several levels of perceived consequentiality and is

[1] Carson and Louviere (2011) distinguish two approaches within stated preference research used in contingent valuation: matching methods, in which respondents provide a number (or numbers) that represent the value of a good to them, and discrete choice experiments, where respondents select their most preferred alternative from a given set of alternatives.

typically located after the preference elicitation task. This method of eliciting the perceptions is widely accepted and applied, but hardly any research has investigated alternative ways of the perception elicitation, with comparing them against this common approach. In this paper, we are specifically interested in understanding the role of the location and possible repetition of the consequentiality perception elicitation question in a stated preference survey. We verify whether stated consequentiality perceptions and stated preferences are affected by placing the consequentiality perception elicitation question before and after preference elicitation tasks.

To that end, we use data from a discrete choice experiment survey that inquires about citizens' preferences towards expansion of urban green spaces in four German cities (Augsburg, Karlsruhe, Leipzig and Nuremberg). The study involves two treatments differing with respect to the consequentiality perception elicitation procedure. In one, respondents are asked about their consequentiality perceptions only after completing all preference elicitation tasks (that is, the perceptions are elicited in a usual way as done in stated preference surveys). In the other, respondents' consequentiality perceptions are elicited twice: before and after the preference elicitation tasks. This design allows us for within-sample and between-sample tests of the effect of the location and repetition of the consequentiality perception elicitation question on stated perceptions and stated preferences.

We are aware of only one study to date that addresses a similar issue. In a stated preference survey concerning drinking water reliability, Lloyd-Smith, Adamowicz and Dupont (forthcoming) ask respondents about consequentiality perceptions either before the preference elicitation task or after it. In all other existing stated preference surveys which collect data on consequentiality perceptions, the question used for the perception disclosure is placed after the preference elicitation part (Lloyd-Smith, Adamowicz and Dupont, forthcoming). Lloyd-Smith, Adamowicz and Dupont (forthcoming) find that the location of consequentiality perception elicitation affects stated consequentiality perceptions: the proportion of respondents who view the survey as inconsequential is by nine percentage points lower among those asked about consequentiality before the preference elicitation than among those asked about it after the preference elicitation (13.3% for the former and 22.3% for the latter).

Our study differs from the research of Lloyd-Smith, Adamowicz and Dupont (forthcoming) in several respects. First, in their survey, consequentiality perceptions are elicited on a five-point Likert scale with only the outer categories labelled (as [the voting results will be] "not taken into account" and "definitely taken into account" [by policy makers], respectively). We use a scale with all categories labelled, ranging from "definitely considered" to "definitely not considered", and including an "I do not know" answer. We note that our research employs a symmetric response scale, as the four labelled responses reflecting defined perceptions of consequentiality (that is, excluding "I do not know" answer) were described using symmetric wording. Second, in their analysis, they use the collected information on perceived consequentiality as a binary variable, where "not taken into account" is coded as one level, while all other categories are combined into a single level expressing some positive degree of consequentiality. We do not convert the ordinal scale responses, and we use the full information as disclosed by respondents. Finally, in contrast to their split-sample design, our study design allows for both within- and between-sample tests whether consequentiality perceptions change between those declared before preference

elicitation and those declared after preference elicitation. Our design also enables verification whether asking about consequentiality perceptions once or twice matters for stated perceptions and stated preferences. An additional difference between our and their study is that they use a single binary choice format to elicit preferences, while we use a sequence of nine three-option choice questions.

Using ordered logit models, we find that stated consequentiality is stronger (i) when respondents are asked about consequentiality perceptions twice (that is, before and after preference elicitation tasks) rather than when asked only once (that is, after the tasks); and (ii) when the perceptions are inquired before the preference elicitation tasks rather than when inquired after the tasks. The latter finding emerges in both within- and between-sample comparisons. The models also suggest that perceived consequentiality does not depend on socio-demographic characteristics of respondents, but rather on the mode of respondents were recruited to the study. Respondents invited to the survey via mail-delivered postcards with a link to the online survey reveal stronger perceived consequentiality than respondents recruited via online panels.

Results of mixed logit willingness-to-pay space models indicate that marginal values for the project attributes significantly differ depending on whether the consequentiality perception elicitation question is asked before the preference elicitation task or is not. Furthermore, the consequentiality perceptions appear to significantly affect stated willingness-to-pay values, and this influence differs depending on whether the consequentiality perception elicitation question is asked before or after preference elicitation tasks. These significant differences in the impact of the way consequentiality perceptions are elicited on willingness-to-pay values for the attributes are observed for each of the four cities in which the survey was conducted.

Our findings suggest that consequentiality perception elicitation may be more complex than it is usually treated in surveys. Both stated consequentiality perceptions and stated preferences could be differently affected by various ways consequentiality perceptions are elicited. Our findings shed some light on how these perceptions change throughout the survey. This evidence also suggests that when willingness-to-pay values are corrected based on stated consequentiality perceptions—as often done in recent stated preference studies—even the corrected values could be sensitive to the way the consequentiality perceptions are elicited. A natural recommendation follows that stated preference researchers should pay enough attention to this piece of a survey structure when designing a survey.

Our study also contributes to the more general literature on directional context effects. In the survey methodology literature (Dillman, 2011; Schuman, Presser and Ludwig, 1981; Tourangeau, Rips and Rasinski, 2000), it is argued that secondary questions in a survey, which are related to the main issue tackled in the survey (in our case, this is the elicitation of preferences towards urban green expansion), provide an interpretative framework that impacts on individual's responses to the main survey questions. Although not broadly investigated, this topic has been undertaken by some studies in the context of preference elicitation surveys. Pouta (2004) shows that asking questions on attitudes and beliefs about the effects of the current and the environmentally-oriented tree cutting practice positively affects willingness-to-pay for forest regeneration revealed in a single binary choice preference elicitation task.

De-Magistris, Gracia and Nayga (2013) test the inclusion of honesty priming tasks before preference elicitation to awake associations of honesty, fairness and truthfulness and, hence, to encourage truthful preference revelation. They observe that the approach works as the willingness-to-pay values from hypothetical preference elicitation questions proceeded by the honesty priming tasks are not significantly different from the values obtained through non-hypothetical (binding) questions. Liebe et al. (2016) find that including questions on anti-Semitic and anti-Arabic attitudes has a positive effect on willingness-to-pay for a "peace product" jointly produced by Israeli and Palestinian producers. They argue that the attitudinal questions lead to an "attention shift". Similarly, asking a question on consequentiality before the preference elicitation can lead to such an attention shift, that is, attention might be directed towards consequences of the responses, thus making the choices in the preference elicitation tasks be viewed as more binding.[2] Our empirical results indicate that asking the consequentiality-perception elicitation question before preference elicitation indeed matters for stated preferences. This effect seems to go in different directions depending on a considered city and an attribute, but the consistent outcome we observe is that on average, asking about consequentiality perceptions before the preference elicitation reduces the disutility from the current state of urban green spaces (that is, the value respondents would need to be paid for keeping the current state). Given that hypothetical choices are typically related to overstated values (Penn and Hu, 2018), our result coincides with the intuition and suggests that asking about perceived consequentiality prior to preference elicitation may work in the expected direction.

The remainder of the paper is structured as follows. Section 2 provides details about the stated preference survey conducted to collect the data. Section 3 outlines the econometric approach used for the data analysis. Section 4 presents the results of the empirical analysis, and Section 5 concludes.

## 2 Empirical data

The survey was conducted in four German cities—Augsburg, Karlsruhe, Leipzig and Nuremberg—as part of the project "Value of Green Urban Spaces: Evaluation, Management and Communication as a key for climate resilient and near-natural green spaces", funded by the German Ministry for Education and Research.[3] The survey questionnaire, implemented online, was split into three parts. In the first part, respondents indicated their place of residence and their most frequently used parks on an embedded map interface. The second part employed a discrete choice experiment to elicit preferences about different forms of expansion of urban green spaces in the respondent's city, with a detailed explanation of the attributes used for the considered project description. This part also included elicitation of

[2] One could argue that asking about perceived consequentiality before preference elicitation can play a similar role as consequentiality scripts (which, at the same time, are observed to generate mixed results regarding their actual influence on respondents' perceptions). However, we highlight the difference: consequentiality scripts might be seen as a "passive" approach, as respondents are only provided with information, while questions eliciting consequentiality perceptions could be viewed as an "active" approach, because respondents need to think about their perceived consequentiality level. This difference in respondents' engagement may impact on that how each of the two approaches work.

[3] More information on the project can be found under the link: https://www.ioew.de/en/project-single/value_of_green_urban_spaces/

respondents' perceptions regarding the survey consequentiality. The third part of the questionnaire held questions on the usage behaviour and attitudes regarding urban green and socio-demographic characteristics.

### 2.1 Discrete choice experiment

The survey considered a city-wide policy project of the extension of urban green. The project was characterised by five attributes, as presented in Table 1, which were explained to respondents in detail on separate screens of the survey prior to the preference elicitation in the discrete choice experiment. All attributes were carefully cross checked with experts from policy and science and a pretest (see section 2.3). The final selection of attributes included the number of street trees per 100 meters, the share of green areas of the city's total area, the share of near-natural green areas of all green areas, the share of pedestrian and cycling greenways[4] of all pedestrian and cycling ways, and a monetary attribute. Each of the four non-monetary attributes took one of three possible levels. One level corresponded to the current city-specific average (that is, a status quo level), and the two other levels represented extensions compared to the current situation, which could be implemented in the city. The status quo levels were derived via GIS data and cross-checked with representatives from the respective city's administration. Respondents were informed about the current average levels in their city for each non-monetary attribute. The levels related to the improvements were displayed in the discrete choice experiment as specific and absolute percentage values, and were not presented as percentage increases as in Table 1 (which are used so for explanation purposes only). The monetary attribute was defined as a compulsory yearly payment per individual that the city would spend exclusively on the development and maintenance of urban green.

Each respondent was provided with a sequence of nine preference elicitation tasks in the discrete choice experiment. Each task included three alternatives, out of which respondents were asked to choose their most preferred alternative. The rightmost alternative was always a status quo described as "Current state", with all attribute levels set to the current averages in a given city. Two other alternatives involved some changes to the current state, and were named as Option 1 and Option 2. Figure 1 shows an example choice task for the city Leipzig.

An orthogonal design with 36 choice tasks split into four blocks was used. A randomly assigned block of nine choice tasks was presented to a respondent. The design was tested via simulation and in the second pilot study. It was compared to various other designs including efficient designs with varying priors and linearity assumptions.

[4] The greenway attribute was described as development of bushes and lawns adjacent to pedestrian and cycling ways to avoid overlap with the street-tree attribute.

*Table 1. Discrete choice experiment attributes and their levels*

| **Attribute** | **Description** | **Levels** | **Status quo levels in the cities** |
|---|---|---|---|
| Street trees | An average number of trees per 100 meters of a street | As today<br>2 trees more per 100 meters<br>4 trees more per 100 meters | Augsburg: 5<br>Karlsruhe: 5<br>Leipzig: 5<br>Nuremberg: 5 |
| Green areas | A share of green spaces in the total area of the city | As today<br>An increase by 5%<br>An increase by 10% | Augsburg: 35%<br>Karlsruhe: 35%<br>Leipzig: 20%<br>Nuremberg: 25% |
| Near-natural green areas | A share of green areas that are near-natural | As today<br>An increase by 10%<br>An increase by 20% | Augsburg: 20%<br>Karlsruhe: 20%<br>Leipzig 20%<br>Nuremberg: 25% |
| Pedestrian and cycling greenways | A share of pedestrian and bike ways that are adjacent to lawns and bushes | As today<br>An increase by 10%<br>An increase by 20% | Augsburg: 25%<br>Karlsruhe: 35%<br>Leipzig: 40%<br>Nuremberg: 35% |
| Cost | A compulsory yearly payment per individual (in EUR) | 0, 5, 10, 50, 100, 200, 300 | |

*Figure 1. An example choice task (for Leipzig)*

| | Option 1 | Option 2 | Current state |
|---|---|---|---|
| **Street trees** | **5 trees** per 100 meters of a street | **9 trees** per 100 meters of a street | **5 trees** per 100 meters of a street |
| **Green areas** | **25%** of the city area is green spaces | **20%** of the city area is green spaces | **20%** of the city area is green spaces |
| **Near-natural green areas** | **30%** of the green areas is near-natural | **40%** of the green areas is near-natural | **20%** of the green areas is near-natural |
| **Pedestrian and cycling greenways** | **60%** of the ways are greenways | **50%** of the ways are greenways | **40%** of the ways are greenways |
| **Cost for you per year** | **€300** | **€100** | **No cost** |
| **Which option do you choose?** | ☐ | ☐ | ☐ |

*Note*: The tasks were originally displayed in German. Here a translation is provided.

## 2.2 Elicitation of consequentiality perceptions

Information about perceived consequentiality was collected in the survey with the use of the question: "To what degree do you believe that your responses will be taken into account in policy and administration?". The Likert response scale included five levels labelled, respectively, as "definitely considered", "rather considered", "rather not considered", "definitely not considered" and "I do not know".[5]

The survey involved two treatments that differed in the way consequentiality perceptions were elicited. In the Asked-Once sample, the consequentiality perception elicitation question was asked, as typically done in valuation surveys, right after the preference elicitation (that is, after the discrete choice experiment). In the Asked-Twice sample, this question appeared in the survey twice: before and after the preference elicitation. Respondents in this sample were not made aware that they would be asked twice about their consequentiality perceptions.

[5] The question and responses were originally in German. Here a translation is provided.

The study discussed in this paper is based on a single question used for consequentiality perception elicitation. In this way, we followed a predominant approach for collecting data on perceived consequentiality as used in stated preference surveys. We note, however, that recent works suggest that consequentiality perceptions could be elicited with the use of a set of questions, in particular, with the use of questions that differentiate between respondents' perceptions towards policy consequentiality and payment consequentiality (e.g., Zawojska, Bartczak and Czajkowski, 2019). Following the definition in Johnston et al. (2017), policy consequentiality can be understood as a positive probability that survey responses will influence decisions related to the outcome in question, while payment consequentiality corresponds to a positive probability that the payment for the considered project will indeed be collected if the project is implemented. Not differentiating between these two aspects of consequentiality, our question grasps to some degree both of them. Understanding how the effect of the location of the consequentiality perception elicitation in a survey differs across these two aspects of consequentiality is beyond the scope of this paper and is left for future work.

## 2.3 Survey implementation

The selection and definition of the attributes, along with the specification of their levels, was guided by the expertise of policy makers from the four cities, and an expert workshop on the value of urban green with policy makers from German cities in March 2018. The development of the questionnaire was consulted through individual interviews with representatives from the general population. Pretesting involved two pilot studies with 192 respondents randomly drawn from the general population of the four cities considered in the research. Input from respondents participating in the pilot studies fostered refinements of the survey, which helped develop the questionnaire for the final data collection.

The final questionnaire was implemented online (Computer-Assisted Web Interviews, CAWI) and consisted of 35 pages. The survey was administered by a professional public opinion polling agency, and the data was collected between July and November 2018.

We received a total of 2,806 completed questionnaires from the four German cities that took part in the survey, including 559 respondents from Augsburg, 479 respondents from Karlsruhe, 1,130 respondents from Leipzig and 638 respondents from Nuremberg. Two modes of recruiting participants to the study were employed. The majority of the sample, amounting to 2,442 respondents, was recruited through online panels of the polling agencies. In order to increase the sample sizes of respondents from Augsburg and Karlsruhe, postcards with a link to the online survey were mailed to 5,500 randomly chosen registered addresses. Out of this group, 228 completed questionnaires from Augsburg and 136 completed questionnaires from Karlsruhe were submitted.[6] Potential duplicates from the different recruiting modes

[6] We do not observe systematic differences in distributions of socio-demographic characteristics across the samples of respondents recruited via online panels and via postcards in Augsburg and Karlsruhe. For both cities, no differences across online- versus mail-recruited samples are found for age, gender and having high-school education. Statistically significant differences emerge only with respect to a household size for Augsburg respondents and with respect to income for Karlsruhe respondents. This evidence of only few differences indicates that the recruitment mode did not critically affect the (self-)selection of respondents.

were eliminated with the use of cookies, filter questions in the questionnaire and individual profiles based on socio-demographic data and location[7].

*Table 2. Socio-demographic characteristics of respondents in the four cities*

| Characteristic | Augsburg | Karlsruhe | Leipzig | Nuremberg |
|---|---|---|---|---|
| **Age** | 46.90 (15.75) | 42.89 (15.49) | 45.01 (15.01) | 47.48 (14.64) |
| **Household size** (including children) | 2.52 (2.21) | 2.56 (2.04) | 2.24 (1.56) | 2.38 (1.61) |
| **Gender** | | | | |
| Female | 51.2% | 49.7% | 52.9% | 52.2% |
| Male | 47.8% | 49.3% | 46.7% | 47.0% |
| **Household monthly income** [EUR] | | | | |
| Less than 1,000 | 5.9% | 8.1% | 8.7% | 5.6% |
| 1,000 - 1,999 | 15.6% | 16.1% | 27.9% | 18.8% |
| 2,000 - 2,999 | 18.2% | 20.5% | 23.6% | 22.3% |
| 3,000 - 3,999 | 20.9% | 15.7% | 12.9% | 13.6% |
| 4,000 - 4,999 | 10.5% | 11.7% | 7.2% | 10.5% |
| 5,000 or more | 9.5% | 11.6% | 3.2% | 8.2% |
| **Education attained** | | | | |
| Secondary or elementary | 25.6% | 21.7% | 23.8% | 37.0% |
| High-school diploma | 33.1% | 32.4% | 37.4% | 29.5% |
| University degree | 37.0% | 43.4% | 36.9% | 29.6% |
| **Number of respondents** | 559 | 479 | 1,130 | 638 |

*Notes*: For Age and Household size, the table shows means (and standard deviations in brackets). For Gender, Household monthly income and Education attained, shares of participants are reported. The shares do not sum up to 100% within cities because of missing observations.

Table 2 shows socio-demographic characteristics of the samples in the four cities. An average respondent in the whole sample was about 46 years old and lived in a two-person household. Out of the total sample, 52% of the respondents are females; 58% of the respondents have a household monthly income of between 1,000 and 3,999 EUR; 27% of the respondents have secondary education, 34% of the

[7] 102 observations were excluded on this basis. The numbers of respondents provided above are after the exclusion of the potential duplicates.

respondents obtained a high-school diploma and 37% of the respondents have a university degree. We note there are some differences in socio-demographic characteristics across the four cities considered, but the differences correspond to differences in the general populations of the cities. In Figure A.1 in the appendix, we present maps with the spatial distribution of respondents in each city.

## 3 Econometric approach

The empirical analysis involves two econometric approaches. To examine how statements about perceived consequentiality are affected by the way the perceptions are elicited, we employ ordered logit models. To inquire whether the effect of consequentiality perceptions on stated preferences differs depending on the way the perceptions are elicited, mixed logit models in willingness-to-pay space are estimated. These two econometric frameworks are described in separate sections below.

In both approaches, information on perceived consequentiality is used. As discussed in Section 2.2, the information was collected on a five-level discrete scale in the survey, with the first four options ordered according to descending belief in the survey consequentiality (to remind, the options were: survey responses were seen that they would be “definitely considered”, “rather considered”, “rather not considered” and “definitely not considered” in policy and administration) and the last option allowing for an indefinite statement (“I do not know”). In all models presented in the paper, we use the collected data on perceived consequentiality as an ordered discrete variable taking four levels, with the level of 1 corresponding to the strongest perceived consequentiality (“definitely considered”) and the level of 4 related to the weakest perceived consequentiality / to the inconsequentiality (“definitely not

considered"). Henceforth, we refer to the variable as *perceived consequentiality*. "I do not know" consequentiality statements are omitted in the econometric modelling.[8,9]

### 3.1 Ordered logit framework

To understand drivers of consequentiality perceptions and, in particular, to examine whether the perceptions are affected by the location and repetition of the consequentiality perception elicitation question, we employ an ordered logit framework (Greene and Hensher, 2010). We use as a dependent variable respondents' statements to the consequentiality perception elicitation question, captured in the variable *perceived consequentiality*, defined as above, while the set of the explanatory variables includes, among others, binary (zero-one-coded) variables controlling for the location and repetition of the consequentiality perception elicitation question (using the notation from the Results section, variables *Before* and *Asked-Twice*, respectively).

Formally, the ordered logit model could be described with the following function:

(1) $y_i^* = x_i^{'}\beta + \varepsilon_i$ ,

where $y_i^*$ is the survey consequentiality in the perception of individual *i*, unobservable directly from a researcher's perspective; $x_i$ is a vector of explanatory variables; $\beta$ is a vector of parameters to be estimated; and $\varepsilon_i$ is the error term with an assumed standard logistic distribution.

[8] This choice is guided by several reasons, which all boil down to selecting a specification that illustrates the problem under investigation in the most straightforward and appropriate way. We estimated and compared model specifications that placed the "I do not know" consequentiality statements in different points on the ordered response scale, namely in the middle of the scale or at one of its ends. The choice of a specification that fits best to the data (based on the Akaike information criterion) was inconclusive: for instance, for the ordered logit models, the model with "I do not know" statements placed as last (least consequential) performed best in the case of the within-sample test, while the model with "I do not know" statements in the middle of the scale resulted in the best fit in between-sample tests. This invoked our further considerations regarding the intuitive interpretation of the scale as potentially viewed by respondents. In the consequentiality perception elicitation question, the "I do not know" option was displayed as last (rightmost), following the "definitely not considered" option. Hence, this could be somewhat unusual that in respondents' cognition, the "I do not know" option took a middle place on the scale. Alternatively, using the "I do not know" responses as actually displayed on the scale in the survey did not appear strongly supported based on the econometric model fitting criteria (such as the Akaike information criterion). What further adds to this confusion is that viewing the "I do not know" response as an even more inconsequential perception than the "definitely not considered" option might seem intuitively inconsistent (although sometimes treated that way in the literature; e.g., Zawojska, Bartczak and Czajkowski, 2019). Taking into account these considerations, we finally decided for excluding "I do not know" responses from the empirical analysis based on the qualitative grounds of the results. The models with "I do not know" responses included revealed the same relationships regarding our crucial variables of interest as the models with "I do not know" responses excluded, with the latter displaying a more clear-cut picture of the relationships.

[9] Appendix Figure A.2 shows the distribution of individuals' consequentiality statements across the five response categories, the option "I do not know" inclusive. Three colours of the columns depict how the distribution differ depending on whether the sample faced the consequentiality-perception elicitation question repeated or not (Asked-Twice vs Asked-Once) and on whether the data comes from the question asked before or after the choice tasks (Before vs After).

Instead of $y_i^*$, we only observe respondents' statements to the consequentiality perception elicitation question, denoted for individual *i* as $y_i$:

$$\begin{aligned} y_i &= 1 \quad \text{if } y_i^* \leq 0 \\ y_i &= 2 \quad \text{if } 0 < y_i^* \leq \mu_2 \\ y_i &= 3 \quad \text{if } \mu_2 < y_i^* \leq \mu_3 \\ y_i &= 4 \quad \text{if } \mu_3 \leq y_i^* \end{aligned} \tag{2}$$

where $\mu$ is a vector of threshold parameters to be estimated.

Then, the probability of observing outcome j, where j = {1,2,3,4}, can be represented as

$$\text{Prob}[y_i = j \mid \mathbf{x}_i] = [F(\mu_j - \boldsymbol{\beta}'\mathbf{x}_i) - F(\mu_{j-1} - \boldsymbol{\beta}'\mathbf{x}_i)] > 0. \tag{3}$$

This leads to the log-likelihood function of a form

$$\log L = \Sigma_{i=1}^{n}\Sigma_{j=0}^{J} m_{ij} \log[F(\mu_j - \boldsymbol{\beta}'\mathbf{x}_i) - F(\mu_{j-1} - \boldsymbol{\beta}'\mathbf{x}_i)], \tag{4}$$

where $m_{ij} = 1$ if $y_i = j$ and 0 otherwise. The vectors of parameters $\beta$ and $\mu$ are estimated with the maximum likelihood method.

### 3.2 Mixed logit willingness-to-pay-space framework

Econometric modelling of preferences disclosed by respondents through discrete choices in preference elicitation tasks in a survey is based on a random utility framework (Hanemann 1984; McFadden 1974). According to the framework, the utility of individual *i* from policy project *p*, $U_{pi}(\cdot)$, consists of two elements: a systematic component and a random component. The systematic component, $V_{pi}(\cdot)$, is a function of observable characteristics of the project, that is, a function of non-monetary attributes used in choice tasks, $\boldsymbol{X}_{pi}$, and of the monetary attribute, cost $C_{pi}$ of the project to the individual. The random component captures factors affecting the utility which cannot be controlled through the observable attributes, and it is treated as an unobservable error term, $\varepsilon_{pi}$. Formally, the utility function can be represented as

$$U_{pi}(\cdot) = V_{pi}(\boldsymbol{X}_{pi}, C_{pi}) + \varepsilon_{pi} = \boldsymbol{\gamma_i}'\boldsymbol{X}_{pi} - \lambda_i C_{pi} + \varepsilon_{pi}, \tag{5}$$

where $\boldsymbol{\gamma_i}$ and $\lambda_i$ are parameters expressing preferences towards non-monetary and monetary policy attributes, respectively. By indexing them over *i*, the parameters are allowed to vary over individuals according to a predefined multivariate distribution (Train, 2009). Such a specification is commonly called a mixed logit approach.

To ease the interpretation and to avoid challenges related to assessing the monetary value of the policy based on a model in preference space with a random coefficient on the cost (Scarpa et al., 2008; Train and Weeks, 2005), we estimate the mixed logit model in willingness-to-pay space. To this end, we define $\lambda_i = {}^{\alpha_i}/_{\mu_i}$ as the cost coefficient in preference space, where $\alpha_i$ stands for the underlying marginal utility

of income and $\mu_i$ is a scale parameter, and $\boldsymbol{\gamma_i} = \boldsymbol{\beta_i}/_{\mu_i}$ as a vector of preference-space coefficients on non-monetary attributes, where $\boldsymbol{\beta_i}$ denotes a vector of underlying marginal utilities associated with these attributes. Further, we assume the error term $\varepsilon_{pi}$ to have an i.i.d. type I extreme value distribution with constant variance $var(\varepsilon_{pi}) = \pi^2/_6$ (Scarpa et al., 2008; Train and Weeks, 2005). Based on (5), the value of marginal willingness-to-pay (an implicit price) for a change in a given non-monetary attribute could be calculated by dividing the coefficient on the attribute by the cost coefficient. Consequently, we can denote a vector of marginal willingness-to-pay values for non-monetary attributes as $\boldsymbol{\omega}_i = \boldsymbol{\gamma}_i/_{\lambda_i} = \boldsymbol{\beta}_i/_{\alpha_i}$. Then, specification (5), defined in preference space, could be reformulated into a behaviorally equivalent specification in willingness-to-pay space:

(6) $U_{pi}(\cdot) = \lambda_i[(\boldsymbol{\gamma_i}/\lambda_i)'\boldsymbol{X}_{pi} - C_{pi}] + \varepsilon_{pi} = \lambda_i(\boldsymbol{\omega}_i'\boldsymbol{X}_{pi} - C_{pi}) + \varepsilon_{pi}.$

Similarly to (5), specification (6) enables the preference parameters in willingness-to-pay space, $\boldsymbol{\omega}_i$, to vary across individuals. We assume these parameters are taken from a normal distribution, and we further specify $\lambda_i = -e^{v_i}$, with $\lambda_i$ being an underlying latent normal factor that defines the lognormal distribution of the cost coefficient. The latter assumption ensures that the marginal utility of income is positive.

The purpose of employing the mixed logit framework is to examine whether the effect of consequentiality perceptions on stated preferences varies as a result of the way the perceptions are elicited. To this end, we extend (6) to allow preferences to vary across respondents' stated perceptions of the survey consequentiality. In that, we follow a general approach as outlined by Czajkowski et al. (2017) and specify the vector of preference parameters in willingness-to-pay space as

(7) $\boldsymbol{\omega}_i = \boldsymbol{\omega}_i^* + \boldsymbol{\rho} S_i,$

where $\boldsymbol{\omega}_i^*$ has a multivariate normal distribution with a set of means and a covariance matrix to be estimated; $\boldsymbol{\rho}$ is a vector of parameters to be estimated; and $S_i$ is the *perceived consequentiality* variable, treated as continuous and normalised to have a zero mean and a unit standard deviation. Hence, the vector $\boldsymbol{\omega}_i^*$ represents willingness-to-pay values for the non-monetary attributes for respondents with a consequentiality perception at the average level. Using the same notation, we also redefine the cost coefficient as

(8) $\lambda_i = -e^{[v_i + \boldsymbol{\tau} S_i]},$

with coefficient τ to be estimated. Equation (8) enables marginal utility of income and scale to vary systematically depending on the level of *perceived consequentiality*.

As the cities involved in the study differed in the current state (status quo) levels of the policy attributes, and, thus, respondents in each city likely valued differently the considered improvements in urban green spaces, we analyse every city separately. We estimate three mixed logit models for each city. Every model uses for the *perceived consequentiality* variable data collected through a different question and/or from a different sample: through the question displayed before choice tasks for the sample asked twice, through the question displayed after choice tasks for the sample asked twice, and through the question displayed after the choice tasks for the sample asked once.

The models are estimated with the maximum simulated likelihood method using 1,000 Sobol draws. Various model specifications were evaluated to assure robustness of our findings. Results from different specifications were consistent with those presented in the next section, with key findings remaining unchanged.

# 4 Results

Organisation of this section reflects our two areas of interest in this study, namely whether stated consequentiality perceptions are driven by the location and possible repetition of the question used for the consequentiality perception elicitation, and whether the effect of consequentiality perceptions on stated preferences differs depending on the way (the location and possible repetition) the perceptions are elicited. Each of these issues is separately addressed in the sections below, using the methodology as described in Section 3.

## 4.1 Drivers of consequentiality perceptions

To formally examine whether the location and repetition of the consequentiality perception elicitation question impacts on stated consequentiality perceptions, we estimate three ordered logit models, in which data for the four cities are pooled. Model 1 provides insight into the role of the location of the question, as it includes respondents from the Asked-Twice sample only and uses as a dependent variable the responses to both consequentiality questions (that is, placed before and after preference elicitation). This way, Model 1 offers a within-sample test whether the location of the question matters for stated consequentiality perceptions. To account for the fact that there are two observations for every individual included in the model, we report robust standard errors clustered by respondents. Model 2 is estimated on data for all respondents, including both treatment samples, and employs as a dependent variable the responses to the consequentiality-perception elicitation question that was asked as first (that is, for Asked-Twice respondents, those are responses from the question displayed before preference elicitation, while for Asked-Once respondents, those are responses from the question displayed after preference elicitation). This model complements Model 1 in examining the role of the location of the question, but by conducting a between-sample test. In this model, every single observation comes from a different respondent. Finally, Model 3 tackles on the issue of the repetition of the consequentiality-perception elicitation question, as it involves all respondents and uses as a dependent variable the responses to the question asked after the preference elicitation. Hence, again, every observation included in this model is of a different respondent.

Each model includes in the set of the explanatory variables either a zero-one-coded variable *Before* equal to 1 when the consequentiality-perception elicitation question was displayed before the choice tasks, or a zero-one-coded variable *Asked-Twice* equal to 1 for respondents who were asked twice about their perceived consequentiality. These two variables allow us to investe the effects of the location and repetition of the consequentiality-perception elicitation question on stated perceptions. In addition to these variables, a set of socio-demographic characteristics are used as control variables in the regressions: a zero-one-coded variable *Male* taking a value of 1 for males, a continuous variable *Age* representing a

respondent's age in years, a group of zero-one-coded variables denoting highest levels of education attained (*Secondary or elementary*, *High-school*, *University*), a group of zero-one-coded variables corresponding to a place of residence and a recruitment mode (*Augsburg Online*, *Augsburg Postal*, *Karlsruhe Online*, *Karlsruhe Postal*, *Leipzig*, where Nuremberg is the reference level).

Table 3 presents the estimation results. Nearly no socio-demographic variables, except for those specific to the recruitment mode, explain variation in stated consequentiality perceptions. The results consistently show for each model specification that the recruitment via postcards strengthens the perceived consequentiality, and the effect is observed for both cities in which some respondents were recruited with this mode (that is, in Augsburg and Karlsruhe).

*Table 3. Ordered logit models of drivers of stated consequentiality perceptions*

| | **Model 1** | **Model 2** | **Model 3** |
|---|---|---|---|
| Dependent variable | Responses to both consequentiality questions | Responses to the consequentiality question asked as first | Responses to the consequentiality question asked after preference elicitation |
| Sample | Asked-Twice | All | All |
| *Before* | -0.246*** (0.049) | --- | --- |
| *Asked-Twice* | --- | -0.764*** (0.082) | -0.493*** (0.081) |
| *Male* | 0.055 (0.104) | -0.029 (0.080) | 0.080 (0.081) |
| *Age* | 0.003 (0.003) | 0.002 (0.003) | 0.003 (0.003) |
| *Secondary or elementary* | -0.061 (0.406) | 0.141 (0.262) | 0.126 (0.270) |
| *High-school* | -0.382 (0.403) | -0.061 (0.259) | -0.163 (0.268) |
| *University* | -0.442 (0.398) | -0.270 (0.258) | -0.228 (0.267) |
| *Augsburg Online* | 0.084 (0.178) | 0.016 (0.142) | 0.128 (0.143) |
| *Augsburg Postal* | -0.515*** (0.196) | -0.526*** (0.160) | -0.496*** (0.160) |
| *Karlsruhe Online* | -0.101 (0.190) | -0.285** (0.141) | -0.222 (0.144) |
| *Karlsruhe Postal* | -0.550** (0.219) | -0.568*** (0.196) | -0.677*** (0.196) |
| *Leipzig* | 0.178 (0.133) | 0.129 (0.105) | 0.106 (0.106) |
| Log-likelihood at convergence | -2,433.2 | -2,509.6 | -2,453 |
| Log-likelihood at constant(s) only | -2,465.5 | -2,582.4 | -2,499 |
| AIC | 4,894.4 | 5,047.2 | 4,933.9 |
| BIC | 4,976 | 5,129.2 | 5,015.9 |
| Number of observations | 2,514 | 2,600 | 2,580 |

*Notes*: ***, **, and * indicate 1%, 5%, and 10% significance levels, respectively. Standard errors are given in brackets (for Model 1 clustered by respondents).

Regarding our research topic, variables *Before* and *Asked-Twice* are of the main interest. In all specifications, the coefficients on these variables are highly significant and negative; that is, given the definition of *perceived consequentiality* employed as the dependent variable, these variables strengthen the consequentiality perception.

Based on Model 1, this implies that when looking at the same respondents, the respondents reveal stronger belief in consequentiality when asked about it before the preference elicitation than when asked after it. From Model 2, we see that when looking at respondents asked about consequentiality before the preference elicitation and those asked after it, the former are characterised by stronger belief in consequentiality. These two models lead to the conclusion that perceived consequentiality might be stronger before facing preference elicitation tasks than after it. This overlaps with the findings of Lloyd-Smith, Adamowicz and Dupont (forthcoming).

In turn, results of Model 3 suggest that respondents asked twice about consequentiality perceptions are characterised by stronger consequentiality belief as reported after the preference elicitation than respondents asked about the perceptions only after the preference elicitation. This could be a signal that perceived consequentiality might be positively influenced with a larger emphasis placed on it through asking questions about perceived consequentiality.

## 4.2 Consequentiality perceptions as a shifter of stated preferences

Results presented in this section aim at answering whether the effect of consequentiality perceptions on stated preferences differs depending on the way the perceptions are elicited; particularly, whether it depends on the location and possible repetition of the consequentiality-perception elicitation question. To that end, and as explained in detail in Section 3, we estimate three mixed logit models in willingness-to-pay space for each city separately.

Given limited space and qualitative similarity of the results, we focus in this section on the sample of Leipzig respondents only. The selection of this group of respondents is based on having the largest number of observations from this city (1,130 respondents). At the end of this section, the results are briefly compared to the other cities to verify robustness of our findings. Detailed results from the other three cities can be found in the Appendix (Tables A.1, A.2 and A.3).

Estimation results of the mixed logit models for Leipzig are provided in Table 4. *Status quo* is a binary variable taking a value of one for the current state and zero otherwise. The remaining variables correspond to the attributes as described in Table 1 and are treated as continuous variables. All preference parameters are modelled as normally-distributed random parameters, except for the cost parameter, which is assumed to follow a log-normal distribution (the estimates of the underlying normal distribution are given). To ensure the model convergence, the attributes *Green areas*, *Near-natural green*, *Greenways* and *Cost* are divided by 100, and we use a negative of the annual cost. Consequently, the means of the preference parameters (denoted with (1)) can be readily interpreted as mean willingness-

to-pay, with the values for the status quo and street trees being expressed in 100 EUR and the values for the remaining non-monetary attributes being in EUR per unit percentage change in the attribute level.

Model I and Model II are estimated on the Asked-Twice sample, while Model III is based on the Asked-Once sample. The means of the preference parameters are in each model interacted with the variable *perceived consequentiality* that has been normalised to have a zero mean and a unit standard deviation. Hence, the means (denoted by (1); that is, when the interactions are ignored) express the mean values for a respondent holding average consequentiality perceptions in the sample. In Model I, the means are interacted with the *perceived consequentiality* as stated in the question preceding the preference elicitation (*Before*). In Model II and Model III, the means are interacted with the *perceived consequentiality* revealed in the question following the preference elicitation (*After*). How the willingness-to-pay values for the attributes change in perceived consequentiality is captured by the coefficients on interactions marked with (3) in Table 4.[10]

On average, respondents are willing to pay annually from 11 to 30 EUR to avoid the current state of the urban green spaces in their city, about 4-7 EUR for an additional tree per 100 meter street, 0.58-1.79 EUR for an additional one percent of green spaces in the total city area, 0.81-0.86 EUR for an additional one percent of near-natural green spaces in the total city green space, and 0.76-1.27 EUR for an additional one percent of greenways in the total pedestrian and bike ways. The means of the preference parameters are significant for all attributes, with the exception of *Status quo* in Model II. All standard deviation parameters are highly significant and, except for *Greenways* and *Cost*, larger than the respective means, which indicates substantial preference heterogeneity. This justifies the use of the mixed logit framework.

Turning into the effect of consequentiality perceptions on stated preferences, we examine the interaction terms. The interaction terms suggest that stated consequentiality perceptions mainly shift respondents' preferences towards the *Status quo*. The direction of the effect is consistent across the three models and indicates that when *perceived consequentiality* gets stronger (negative values of the *perceived consequentiality* variable), the disutility from the current state is intensified. In other words, respondents with strong consequentiality perceptions are willing to pay more to avoid the current state than respondents with weak consequentiality perceptions, where the latter might be even willing to pay to keep the current state. With respect to the willingness-to-pay values for the remaining attributes, the effect of consequentiality perceptions differs across the model specifications. When the consequentiality statements from the *After* question are used, stronger perceived consequentiality increases willingness-to-pay for *Green areas* and *Near-natural green* among the Asked-Twice respondents (Model II), while the

[10] Although Models I and II are estimated on the same samples, we note that the estimates of the means and standard deviations of preference parameters may differ between the two specifications, as the *perceived consequentiality* variable employed in the interactions in each of them comes from different questions (*Before* or *After* the choice tasks). Although this variable is normalised to have a zero mean and a unit standard deviation, the measurement of consequentiality through such responses can be imprecise and subject to measurement error. Because of this possible within- and between-individual variation in interpreting the possible responses on the Likert scale, the differences in the means and standard deviations of preference parameters are justified.

effect is opposite among the Asked-Once respondents (Model I).[11] The remaining interaction terms are not statistically significant.

The question of our main interest is whether the effect of consequentiality perceptions on the willingness-to-pay values differs depending on the way (the location and possible repetition) the perceptions are elicited. To that end, we formally test for statistical differences in the coefficients on the means and the interaction terms across the three models using z-tests. We compare each pair of the coefficients and report the results in Table 5.

*Table 4. Mixed logit models in willingness-to-pay space for Leipzig*

| | **Model I** | **Model II** | **Model III** |
|---|---|---|---|
| Sample | Asked-Twice | Asked-Twice | Asked-Once |
| Means interacted with | *Before* | *After* | *After* |
| Means | | | |
| *Status quo* (1) | -0.11 (0.03)*** | -0.01 (0.03) | -0.30 (0.03)*** |
| *Street trees* (1) | 0.07 (0.01)*** | 0.07 (0.01)*** | 0.04 (0.01)*** |
| *Green areas* (1) | 1.79 (0.29)*** | 1.58 (0.27)*** | 0.58 (0.23)** |
| *Near-natural green* (1) | 0.84 (0.14)*** | 0.81 (0.12)*** | 0.86 (0.10)*** |
| *Greenways* (1) | 1.27 (0.11)*** | 1.19 (0.12)*** | 0.76 (0.10)*** |
| A negative of *Cost* (1) | 1.57 (0.15)*** | 1.43 (0.10)*** | 1.58 (0.12)*** |
| Standard Deviations | | | |
| *Status quo* (2) | 1.62 (0.06)*** | 1.63 (0.05)*** | 1.57 (0.04)*** |
| *Street trees* (2) | 0.09 (0.01)*** | 0.11 (0.01)*** | 0.07 (0.00)*** |
| *Green areas* (2) | 2.46 (0.51)*** | 2.38 (0.26)*** | 1.38 (0.23)*** |
| *Near-natural green* (2) | 1.98 (0.19)*** | 1.45 (0.13)*** | 1.40 (0.11)*** |
| *Greenways* (2) | 0.74 (0.16)*** | 0.71 (0.10)*** | 0.40 (0.09)*** |
| A negative of *Cost* (2) | 1.46 (0.17)*** | 1.35 (0.11)*** | 1.55 (0.13)*** |
| Interactions with *perceived consequentiality* | | | |
| *Status quo* (3) | 0.20 (0.02)*** | 0.08 (0.03)** | 0.21 (0.03)*** |
| *Street trees* (3) | 0.00 (0.01) | -0.01 (0.01) | 0.00 (0.01) |
| *Green areas* (3) | -0.13 (0.20) | -0.78 (0.28)*** | 0.34 (0.21)* |
| *Near-natural green* (3) | -0.08 (0.12) | -0.27 (0.13)** | 0.24 (0.12)** |
| *Greenways* (3) | -0.16 (0.10) | 0.09 (0.12) | -0.08 (0.10) |
| A negative of *Cost* (3) | 0.10 (0.09) | 0.11 (0.08) | 0.02 (0.09) |

[11] Although this may appear surprising that the effect of consequentiality perceptions on willingness-to-pay goes in opposite directions for different samples, we note that similar findings of opposing effects are identified in other valuation studies (Zawojska, Bartczak and Czajkowski, 2019, and references therein). This could be related to the fact that it is difficult to assess what is revealed by fully hypothetical respondents (Carson and Groves 2007), and so the effect could go either way.

| | | | |
|---|---|---|---|
| Log-likelihood at convergence | -2,995.72 | -2,992.61 | -2,925.50 |
| Log-likelihood at constant(s) only | -5,013.59 | -4,953.57 | -4,875.63 |
| McFadden's pseudo-$R^2$ | 0.40 | 0.40 | 0.40 |
| Ben-Akiva-Lerman's pseudo-$R^2$ | 0.57 | 0.56 | 0.56 |
| AIC | 6,027.44 | 6,021.21 | 5,887.00 |
| BIC | 6,144.04 | 6,137.44 | 6,002.77 |
| Number of observations | 4,806 | 4,707 | 4,590 |
| Number of respondents | 534 | 523 | 510 |
| Number of parameters | 18 | 18 | 18 |

*Notes*: ***, **, and * indicate 1%, 5%, and 10% significance levels, respectively. Standard errors are given in brackets.

*Table 5. Differences in coefficients and Z-test results of significance of the differences between the interaction-term coefficients for Leipzig*

| | **$H_0$: Model I – Model II = 0** | | | | **$H_0$: Model I – Model III = 0** | | | | **$H_0$: Model II – Model III = 0** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Means** | | **Interactions (*perc. cons.*)** | | **Means** | | **Interactions (*perc. cons.*)** | | **Means** | | **Interactions (*perc. cons.*)** | |
| *Status quo* | -0.095 | ** | 0.119 | *** | 0.196 | *** | -0.015 | | 0.291 | *** | -0.135 | *** |
| *Street trees* | 0.002 | | 0.012 | | 0.026 | *** | -0.002 | | 0.025 | *** | -0.014 | |
| *Green areas* | 0.206 | | 0.647 | * | 1.201 | *** | -0.473 | * | 0.995 | *** | -1.120 | *** |
| *Near-natural green* | 0.031 | | 0.188 | | -0.018 | | -0.322 | * | -0.048 | | -0.511 | *** |
| *Greenways* | 0.084 | | -0.244 | | 0.510 | *** | -0.081 | | 0.426 | *** | 0.163 | |

*Notes*: The numbers inform on values of the differences between respective coefficients. The differences were calculated as indicated in the first row of the table (taking coefficient values from respective models), and $H_0$ denotes the null hypothesis tested. *perc. cons.* is an abbreviation for *perceived consequentiality*. ***, **, and * indicate 1%, 5%, and 10% significance levels, respectively.

The comparison in the first column disentangles the effect of the location of the consequentiality-perception elicitation question, based on a within-sample verification. Both Model I and Model II are estimated on the same sample of respondents (Asked-Twice), but use, respectively, *Before* and *After* consequentiality statements for the interaction term. Except for the *Status quo*, we do not find significant differences in mean willingness-to-pay, which aligns with the expectation as the two models employ the same sample of respondents. The significant difference in the value of the *Status quo* is comparatively small, amounting to 0.09. Nevertheless, we observe some differences in the impact of *perceived consequentiality* on stated values. The statistically significant differences emerge with respect to the effect on the willingness-to-pay for *Status quo* and *Green areas*.

The next comparison concerns Model I and Model III. This comparison also adds to the effect of the location of the consequentiality-perception elicitation question, but based on a between-sample verification. Here, we find that the mean willingness-to-pay values for nearly all of the attributes are significantly different across the two model specifications. On average, and when controlled for perceived consequentiality, the mean values declared by respondents who faced the consequentiality-perception elicitation question before the choice tasks (the Asked-Twice sample) are higher than the mean values declared by respondents who were not asked this question before the choice tasks (the Asked-Once sample). Hence, this result could imply that asking a consequentiality-perception elicitation question before the preference elicitation positively influences willingness-to-pay, ceteris paribus. The effect of perceived consequentiality as stated in the first question seen by the respondents on the willingness-to-pay values is only slightly different between the two models—only at 10% significance for the attributes *Green areas* and *Near-natural green.*

The final column allows us to test the role of a repetition of the consequentiality perception elicitation question on stated preferences. With respect to the mean values of the attributes, similarly significant differences are found, as in the previous comparison, which is consistent as both comparisons employ the same samples of respondents. However, the comparison of Model II and Model III reveals some significant differences in the impact of *perceived consequentiality* on willingness-to-pay values. As mentioned earlier, for the attributes *Green areas* and *Near-natural green*, the effect of *perceived consequentiality* on stated willingness-to-pay values appears to go in opposite directions depending on whether consequentiality statements from a repeated or a single question are used. Specifically, when the *perceived consequentiality* is measured through a repeated question, respondents' willingness-to-pay seems to increase with the strength of the consequentiality perception. Instead, when the *perceived consequentiality* is measured through a single question, their willingness-to-pay appears to decrease with the increasing degree of consequentiality belief. This result emphasises that the perceived consequentiality could importantly vary in the impact on stated preferences depending on how the perceptions are elicited.

The estimation results for the other cities involved in the study look very similar and are presented in the Appendix Tables A.1, A.2 and A.3. Preference parameters and standard deviations are generally highly statistically significant, with mean willingness-to-pay values for the current state (*Status quo*) being negative and mean willingness-to-pay values for non-monetary attributes being positive. Coefficients on interactions with the *perceived consequentiality* differ in significance, but again the consistent result (except for one model) is that the willingness-to-pay value for *Status quo* significantly changes across different strength levels of consequentiality perception. This preference parameter seems to be most consistently affected by perceived consequentiality.

Regarding the z-tests of significance of differences across preference parameters and interactions coefficients across models based on different data for the *perceived consequentiality* variable, similarly to the discussed case of Leipzig, we find many significant differences between the coefficients, pointing to the important impact of the way consequentiality perceptions are elicited on stated preferences. The results are provided in the Appendix Tables A.4, A.5 and A.6.

## 5 Discussion and conclusions

In this paper, we investigated the effect of the location and repetition of a consequentiality elicitation question on the perception of consequentiality and on preference elicitation expressed in terms of willingness-to-pay. Using a split-sample approach we compared responses from a questionnaire displaying a consequentiality perception elicitation question before and after the preference elicitation task with responses from a questionnaire displaying the same question only once after the preference elicitation task. Using ordered logit models, our results indicate that responses depend on both the position of the question and whether the question was asked twice. In the question before the preference elicitation task respondents state higher perceived consequentiality than in the question after the task in both samples. Respondents who answered the consequentiality question twice state higher perceived consequentiality in the question after the choice task than respondents who answered only one question after the choice task. One reason for this observation may be that respondents try to be consistent with their answer to the first question.

Using mixed logit models in willingness-to-pay space, we found that willingness-to-pay increases when the consequentiality question is asked before the preference elicitation task. Additionally, all models showed that a higher stated consequentiality perception leads to a lower likelihood of choosing the status quo alternative. Regarding the effect of perceived consequentiality on willingness-to-pay for attributes, the results are mixed, and seem to depend on the location of the consequentiality question.

Our results also contribute to the relatively small literature on directional context effects. It has been shown that asking context questions about the good to be valued in a preference elicitation task influence (increase) the estimated benefits from the good (Liebe et al., 2016; Pouta, 2004). Similarly, our results indicate the asking a consequentiality question before the choice task increases willingness-to-pay—independent of the respondents' stated consequentiality perception. This result was not confirmed in all city samples, but can have important implications. While the literature consistently states that consequentiality reminders lead to lower willingness-to-pay, our results partially imply the opposite. Putting respondents into a consequentiality context lead to higher willingness-to-pay. Further research at this end could help to better understand directional context effects with respect to consequentiality. The result also shows that willingness-to-pay is sensitive to the context questions asked before the preference elicitation task.

Our study is limited in several aspects. First, our analysis did not include respondents who stated that they "do not know" how they perceive consequentiality. Although only a minority (<10% in all questions) of respondents have selected this option, it may still be an important determinant of preferences. Second, we included the consequentiality responses on a continuous scale, implying linear effects of increasing consequentiality perceptions. Accounting for nonlinearities may provide further insights. For example, Lloyd-Smith, Adamowicz and Dupont (forthcoming) used a binary indicator on whether the strongest consequentiality category has been selected or not. Such a coding was not possible in our sample as only a few people have chosen the outer categories (responses definitely considered, responses definitely not considered). Third, we used only one question on consequentiality. The current literature suggests to use

more than one question to infer consequentiality, especially to distinguish between policy and payment consequentiality. A separate question on payment consequentiality may provide different results. Here, further research on the effect of the type of question could provide further insights on the placement of consequentiality questions.

**Acknowledgements**: Ewa Zawojska gratefully acknowledges the support of the National Science Centre in Poland (Preludium 8, grant no. 2014/15/N/HS4/01328) and the support of the Foundation for Polish Science (FNP).

## Appendix

*Figure A.1 Location of the residence of respondents in the four cities*

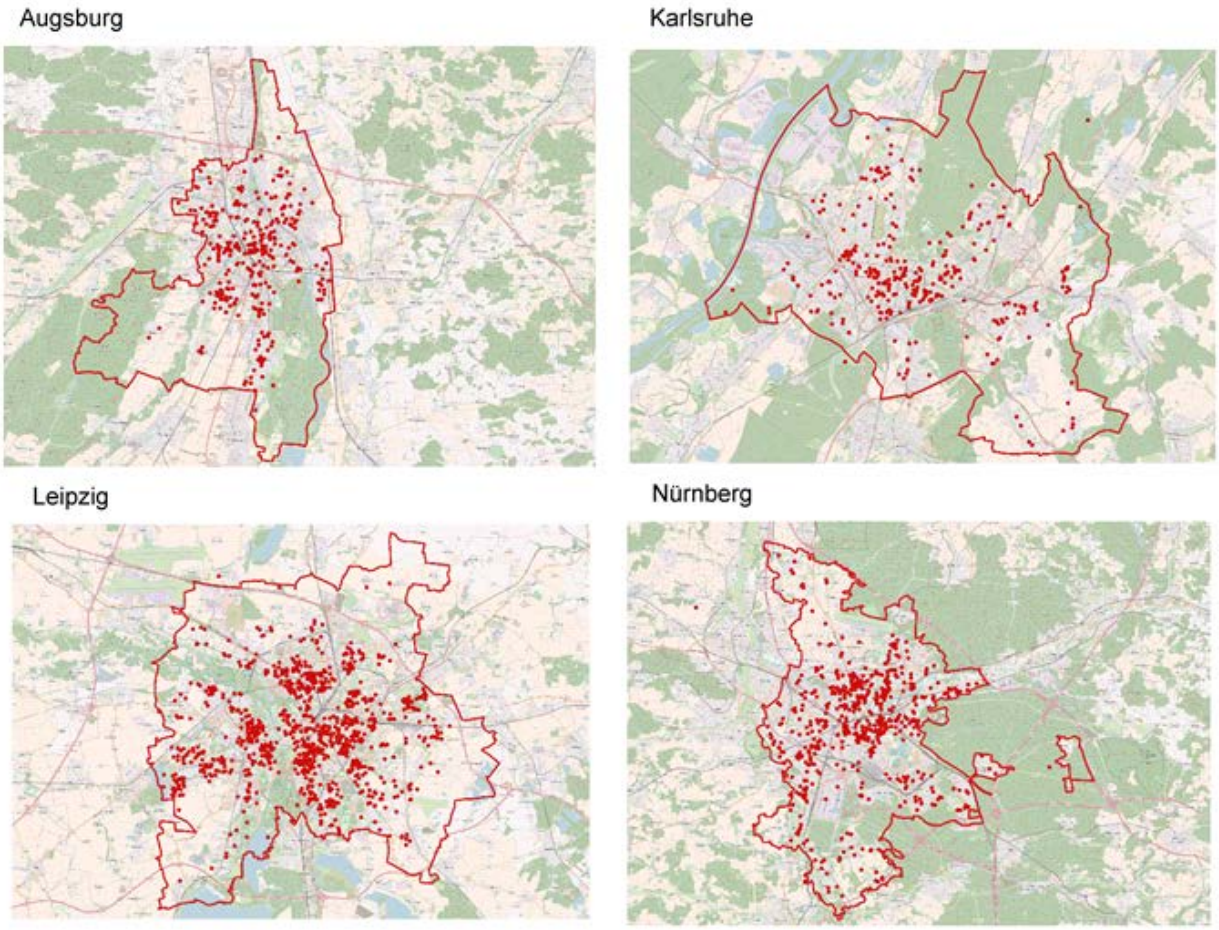

*Figure A.2 Distribution of responses to the consequentiality-perception elicitation question*

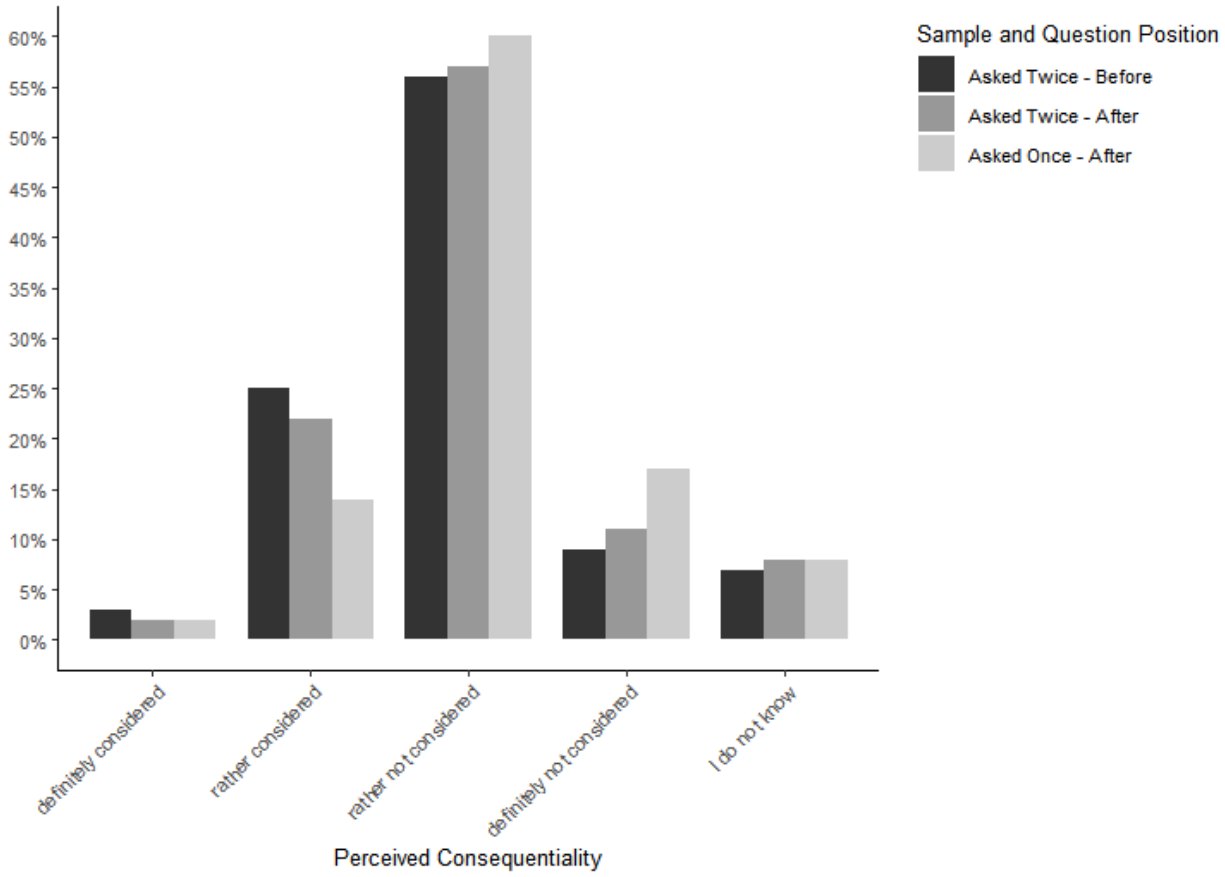

*Table A.1. Mixed logit models in willingness-to-pay space for Augsburg*

| | **Model I** | **Model II** | **Model III** |
|---|---|---|---|
| Sample | Asked-Twice | Asked-Twice | Asked-Once |
| Means interacted with | *Before* | *After* | *After* |
| Means | | | |
| *Status Quo (1)* | -0.58 (0.06)*** | -0.61 (0.05)*** | -0.65 (0.05)*** |
| *Street trees (1)* | 0.09 (0.01)*** | 0.07 (0.01)*** | 0.07 (0.01)*** |
| *Green areas (1)* | 1.63 (0.50)*** | 1.50 (0.46)*** | 0.88 (0.35)** |
| *Near-natural green (1)* | 1.65 (0.27)*** | 1.87 (0.22)*** | 1.35 (0.18)*** |
| *Greenways (1)* | 1.36 (0.20)*** | 1.38 (0.19)*** | 1.06 (0.18)*** |
| *A negative of Cost (1)* | 1.14 (0.12)*** | 1.23 (0.15)*** | 1.33 (0.13)*** |
| Standard Deviations | | | |
| *Status Quo (2)* | 1.60 (0.09)*** | 1.45 (0.10)*** | 1.40 (0.06)*** |
| *Street trees (2)* | 0.13 (0.01)*** | 0.09 (0.01)*** | 0.10 (0.01)*** |
| *Green areas (2)* | 3.63 (0.45)*** | 5.56 (0.56)*** | 1.75 (0.37)*** |
| *Near-natural green (2)* | 3.08 (0.32)*** | 3.30 (0.22)*** | 2.19 (0.20)*** |
| *Greenways (2)* | 0.98 (0.21)*** | 0.67 (0.16)*** | 0.72 (0.21)*** |
| *A negative of Cost (2)* | 1.14 (0.14)*** | 1.23 (0.17)*** | 1.29 (0.14)*** |
| Interactions with perceived consequentiality | | | |
| *Status Quo (3)* | 0.14 (0.06)** | 0.26 (0.04)*** | -0.21 (0.03)*** |
| *Street trees (3)* | 0.01 (0.01) | -0.03 (0.01)*** | -0.01 (0.01) |
| *Green areas (3)* | 0.00 (0.47) | -0.97 (0.50)** | -0.85 (0.27)*** |
| *Near-natural green (3)* | -0.03 (0.26) | -0.32 (0.19)* | -0.13 (0.17) |
| *Greenways (3)* | 0.09 (0.18) | -0.19 (0.16) | 0.26 (0.13)* |
| Log-likelihood at convergence | -1627.34 | -1629.31 | -1591.21 |
| Log-likelihood at constant(s) only | -2571.61 | -2594.38 | -2557.66 |
| McFadden's pseudo-$R^2$ | 0.37 | 0.37 | 0.38 |
| Ben-Akiva-Lerman's pseudo-$R^2$ | 0.52 | 0.53 | 0.53 |
| AIC | 3290.68 | 3294.62 | 3218.43 |
| BIC | 3394.39 | 3398.47 | 3322.07 |
| Number of observations | 2349 | 2367 | 2340 |
| Number of respondents | 261 | 263 | 260 |
| Number of parameters | 18 | 18 | 18 |

Notes: ***, **, and * indicate 1%, 5%, and 10% significance levels, respectively. Standard errors are given in brackets.

*Table A.2. Mixed logit models in willingness-to-pay space for Karlsruhe*

| | **Model I** | **Model II** | **Model III** |
|---|---|---|---|
| Sample | Asked-Twice | Asked-Twice | Asked-Once |
| Means interacted with | *Before* | *After* | *After* |
| Means | | | |
| *Status Quo (1)* | -0.37 (0.07)*** | -0.28 (0.06)*** | -0.69 (0.05)*** |
| *Street trees (1)* | 0.09 (0.01)*** | 0.08 (0.01)*** | 0.08 (0.01)*** |
| *Green areas (1)* | 0.66 (0.50) | 0.99 (0.53)* | 1.77 (0.35)*** |
| *Near-natural green (1)* | 1.07 (0.23)*** | 1.45 (0.30)*** | 1.74 (0.14)*** |
| *Greenways (1)* | 1.37 (0.22)*** | 1.15 (0.23)*** | 0.99 (0.11)*** |
| *A negative of Cost (1)* | 1.25 (0.15)*** | 1.24 (0.15)*** | 1.62 (0.19)*** |
| Standard Deviations | | | |
| *Status Quo (2)* | 1.66 (0.11)*** | 1.79 (0.17)*** | 1.55 (0.07)*** |
| *Street trees (2)* | 0.11 (0.01)*** | 0.12 (0.02)*** | 0.11 (0.01)*** |
| *Green areas (2)* | 1.59 (0.52)*** | 0.19 (0.57) | 1.47 (0.23)*** |
| *Near-natural green (2)* | 2.77 (0.23)*** | 2.59 (0.29)*** | 2.26 (0.12)*** |
| *Greenways (2)* | 1.34 (0.28)*** | 1.11 (0.23)*** | 0.93 (0.12)*** |
| *A negative of Cost (2)* | 1.24 (0.16)*** | 1.25 (0.16)*** | 1.61 (0.22)*** |
| Interactions with perceived consequentiality | | | |
| *Status Quo (3)* | -0.06 (0.07) | 0.27 (0.12)** | 0.15 (0.05)*** |
| *Street trees (3)* | -0.01 (0.01) | -0.00 (0.02) | -0.02 (0.01)* |
| *Green areas (3)* | -0.22 (0.54) | 0.42 (0.64) | 0.24 (0.37) |
| *Near-natural green (3)* | -0.50 (0.16)*** | -0.24 (0.43) | -0.18 (0.13) |
| *Greenways (3)* | -0.17 (0.23) | -0.29 (0.28) | -0.08 (0.12) |
| Log-likelihood at convergence | -1285.76 | -1243.45 | -1340.35 |
| Log-likelihood at constant(s) only | -2111.03 | -2054.13 | -2171.54 |
| McFadden's pseudo-$R^2$ | 0.39 | 0.39 | 0.38 |
| Ben-Akiva-Lerman's pseudo-$R^2$ | 0.55 | 0.55 | 0.53 |
| AIC | 2607.53 | 2522.91 | 2716.70 |
| BIC | 2708.08 | 2622.96 | 2817.33 |
| Number of observations | 1971 | 1917 | 1980 |
| Number of respondents | 219 | 213 | 220 |
| Number of parameters | 18 | 18 | 18 |

Notes: ***, **, and * indicate 1%, 5%, and 10% significance levels, respectively. Standard errors are given in brackets.

*Table A.3. Mixed logit models in willingness-to-pay space for Nuremberg*

| | **Model I** | **Model II** | **Model III** |
|---|---|---|---|
| Sample | Asked-Twice | Asked-Twice | Asked-Once |
| Means interacted with | *Before* | *After* | *After* |
| Means | | | |
| *Status Quo (1)* | -0.22 (0.04)*** | -0.58 (0.03)*** | -0.42 (0.05)*** |
| *Street trees (1)* | 0.08 (0.01)*** | 0.08 (0.01)*** | 0.08 (0.01)*** |
| *Green areas (1)* | 2.78 (0.28)*** | 1.76 (0.31)*** | 1.78 (0.39)*** |
| *Near-natural green (1)* | 0.84 (0.13)*** | 0.86 (0.11)*** | 0.53 (0.17)*** |
| *Greenways (1)* | 0.77 (0.13)*** | 0.75 (0.09)*** | 1.18 (0.18)*** |
| *A negative of Cost (1)* | 1.47 (0.16)*** | 1.57 (0.17)*** | 1.30 (0.12)*** |
| Standard Deviations | | | |
| *Status Quo (2)* | 2.07 (0.10)*** | 1.76 (0.03)*** | 1.65 (0.08)*** |
| *Street trees (2)* | 0.09 (0.01)*** | 0.08 (0.00)*** | 0.09 (0.01)*** |
| *Green areas (2)* | 3.65 (0.14)*** | 3.45 (0.27)*** | 2.92 (0.32)*** |
| *Near-natural green (2)* | 1.15 (0.14)*** | 1.60 (0.08)*** | 1.53 (0.23)*** |
| *Greenways (2)* | 0.86 (0.12)*** | 1.49 (0.06)*** | 0.76 (0.14)*** |
| *A negative of Cost (2)* | 1.61 (0.17)*** | 1.69 (0.19)*** | 1.19 (0.13)*** |
| Interactions with perceived consequentiality | | | |
| *Status Quo (3)* | 0.22 (0.04)*** | 0.31 (0.03)*** | 0.19 (0.05)*** |
| *Street trees (3)* | 0.01 (0.01) | -0.00 (0.01) | 0.01 (0.01)* |
| *Green areas (3)* | 1.00 (0.32)*** | -0.58 (0.29)** | 0.03 (0.31) |
| *Near-natural green (3)* | -0.32 (0.19)* | -0.42 (0.10)*** | 0.04 (0.14) |
| *Greenways (3)* | -0.36 (0.15)** | -0.40 (0.10)*** | 0.18 (0.13) |
| Log-likelihood at convergence | -1759.31 | -1723.73 | -1785.10 |
| Log-likelihood at constant(s) only | -2865.40 | -2820.97 | -2907.46 |
| McFadden's pseudo-$R^2$ | 0.39 | 0.39 | 0.39 |
| Ben-Akiva-Lerman's pseudo-$R^2$ | 0.55 | 0.55 | 0.54 |
| AIC | 3554.61 | 3483.45 | 3606.20 |
| BIC | 3660.59 | 3589.12 | 3712.42 |
| Number of observations | 2664 | 2619 | 2700 |
| Number of respondents | 296 | 291 | 300 |
| Number of parameters | 18 | 18 | 18 |

Notes: ***, **, and * indicate 1%, 5%, and 10% significance levels, respectively. Standard errors are given in brackets.

*Table A.4. Differences in coefficients and Z-test results of significance of the differences between the interaction-term coefficients for Augsburg*

| | **$H_0$: Model I – Model II = 0** | | | **$H_0$: Model I – Model III = 0** | | | **$H_0$: Model II – Model III = 0** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Means** | **Interactions (*perc. cons.*)** | | **Means** | **Interactions (*perc. cons.*)** | | **Means** | **Interactions (*perc. cons.*)** | |
| *Status quo* | 0.0353 | -0.1225 | * | 0.0658 | 0.3501 | *** | 0.0306 | 0.4725 | *** |
| *Street trees* | 0.0151 | 0.0364 | ** | 0.0163 | 0.0160 | | 0.0013 | -0.0205 | * |
| *Green areas* | 0.1292 | 0.9756 | | 0.7486 | 0.8537 | | 0.6194 | -0.1219 | |
| *Near-natural green* | -0.2210 | 0.2976 | | 0.3033 | 0.1053 | | 0.5243 | -0.1923 | * |
| *Greenways* | -0.0248 | 0.2826 | | 0.2983 | -0.1697 | | 0.3231 | -0.4523 | ** |

*Note*: The same notes apply as to Table 5.

*Table A.5. Differences in coefficients and Z-test results of significance of the differences between the interaction-term coefficients for Karlsruhe*

| | **$H_0$: Model I – Model II = 0** | | | **$H_0$: Model I – Model III = 0** | | | **$H_0$: Model II – Model III = 0** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Means** | **Interactions (*perc. cons.*)** | | **Means** | | **Means** | **Interactions (*perc. cons.*)** | | **Means** |
| *Status quo* | -0.0952 | -0.3230 | ** | 0.3186 | *** | -0.2102 | *** | 0.4138 | *** | 0.1128 |
| *Street trees* | 0.0082 | -0.0039 | | 0.0138 | | 0.0094 | | 0.0057 | | 0.0134 |
| *Green areas* | -0.3329 | -0.6424 | | -1.1124 | * | -0.4630 | | -0.7795 | | 0.1794 |
| *Near-natural green* | -0.3770 | -0.2561 | | -0.6678 | ** | -0.3221 | | -0.2908 | | -0.0661 |
| *Greenways* | 0.2175 | 0.1214 | | 0.3721 | | -0.0929 | | 0.1546 | | -0.2144 |

*Note*: The same notes apply as to Table 5.

*Table A.6. Differences in coefficients and Z-test results of significance of the differences between the interaction-term coefficients for Nuremberg*

| | **$H_0$: Model I – Model II = 0** | | | | **$H_0$: Model I – Model III = 0** | | | | **$H_0$: Model II – Model III = 0** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Means** | | **Interactions (*perc. cons.*)** | | **Means** | | **Means** | | **Interactions (*perc. cons.*)** | | **Means** | |
| *Status quo* | 0.3564 | *** | -0.0891 | | 0.1922 | *** | 0.0321 | | -0.1642 | *** | 0.1211 | ** |
| *Street trees* | 0.0013 | | 0.0100 | | 0.0026 | | -0.0018 | | 0.0013 | | -0.0118 | |
| *Green areas* | 1.0232 | ** | 1.5788 | *** | 0.9982 | ** | 0.9740 | ** | -0.0249 | | -0.6048 | |
| *Near-natural green* | -0.0183 | | 0.0998 | | 0.3084 | | -0.3639 | | 0.3267 | | -0.4637 | *** |
| *Greenways* | 0.0167 | | 0.0441 | | -0.4167 | * | -0.5386 | *** | -0.4334 | ** | -0.5827 | *** |

*Note*: The same notes apply as to Table 5.

## References


Andor, M. A., Frondel, M., & Vance, C. (2017). Mitigating Hypothetical Bias: Evidence on the Effects of Correctives from a Large Field Study. *Environmental and Resource Economics*, *68*(3), 777-796.

Broadbent, C. D. (2012). Hypothetical bias, consequentiality and choice experiments. *Economics Bulletin*, *32*(3), 2490-2499.

Broadbent, C. D., Grandy, J. B., & Berrens, R. P. (2010). Testing for hypothetical bias in a choice experiment using a local public good: riparian forest restoration. *International Journal of Ecological Economics and Statistics*, *19*(F10), 1-19.

Carson, R. T., & Groves, T. (2007). Incentive and informational properties of preference questions. *Environmental and resource economics*, *37*(1), 181-210.

Carson, R. T., Groves, T., & List, J. A. (2014). Consequentiality: A theoretical and experimental exploration of a single binary choice. *Journal of the Association of Environmental and Resource Economists*, *1*(1), 171-207.

Carson, R. T., & Louviere, J. J. (2011). A common nomenclature for stated preference elicitation approaches. *Environmental and Resource Economics*, *49*(4), 539-559.

Czajkowski, M., Vossler, C. A., Budziński, W., Wiśniewska, A., & Zawojska, E. (2017). Addressing empirical challenges related to the incentive compatibility of stated preferences methods. *Journal of Economic Behavior & Organization*, *142*, 47-63.

De-Magistris, T., Gracia, A., & Nayga Jr, R. M. (2013). On the use of honesty priming tasks to mitigate hypothetical bias in choice experiments. *American Journal of Agricultural Economics*, 95(5), 1136-1154.

Dillman, D. A. (2011). *Mail and Internet surveys: The tailored design method – 2007 Update with new Internet, visual, and mixed-mode guide*. John Wiley & Sons.

Drichoutis, A. C., Vassilopoulos, A., Lusk, J. L., & Nayga, R. M. (2017). Consumer preferences for fair labour certification. *European Review of Agricultural Economics*, *44*(3), 455-474.

Greene, W. H., & Hensher, D. A. (2010). *Modeling ordered choices: A primer*. Cambridge University Press.

Groothuis, P. A., Mohr, T. M., Whitehead, J. C., & Cockerill, K. (2017). Endogenous Consequentiality in Stated Preference Referendum Data: The Influence of the Randomly Assigned Tax Amount. *Land Economics*, *93*(2), 258-268.

Hanemann, W. M. (1984). Welfare evaluations in contingent valuation experiments with discrete responses. *American journal of agricultural economics*, *66*(3), 332-341.

Herriges, J., Kling, C., Liu, C. C., & Tobias, J. (2010). What are the consequences of consequentiality?. *Journal of Environmental Economics and Management*, *59*(1), 67-81.

Interis, M. G., & Petrolia, D. R. (2014). The Effects of Consequentiality in Binary-and Multinomial-Choice Surveys. *Journal of Agricultural and Resource Economics*, 201-216.

Johnston, R. J., Boyle, K. J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T. A., Hanemann, W. M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R. & Vossler, C. A. (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, *4*(2), 319-405.

Liebe, U., Hundeshagen, C., Beyer, H., & von Cramon-Taubadel, S. (2016). Context effects and the temporal stability of stated preferences. *Social science research*, *60*, 135-147.

Lloyd-Smith, P., Adamowicz, W., & Dupont, D., (forthcoming). Incorporating stated consequentiality questions in stated preference research. *Land Economics*.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In: Zarembka, P. (Ed.) *Frontiers in econometrics* (pp. 105–142). Academic Press.

Oehlmann, M., & Meyerhoff, J. (2017). Stated preferences towards renewable energy alternatives in Germany–do the consequentiality of the survey and trust in institutions matter? *Journal of Environmental Economics and Policy*, *6*(1), 1-16.

Penn, J. M., & Hu, W. (2018). Understanding hypothetical bias: An enhanced meta-analysis. *American Journal of Agricultural Economics, 100*(4), 1186-1206

Pouta, E. (2004). Attitude and belief questions as a source of context effect in a contingent valuation survey. *Journal of Economic Psychology*, *25*(2), 229-242.

Scarpa, R., & Rose, J. M. (2008). Design efficiency for non-market valuation with choice modelling: how to measure it, what to report and why. *Australian journal of agricultural and resource economics*, *52*(3), 253-282.

Schuman, H., Presser, S., & Ludwig, J. (1981). Context effects on survey responses to questions about abortion. *Public Opinion Quarterly*, *45*(2), 216-223.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.

Train, K., & Weeks, M. (2005). Discrete choice models in preference space and willingness-to-pay space. In Scarpa, R., & Alberini, A. (Eds.) *Applications of simulation methods in environmental and resource economics* (pp. 1-16). Springer, Dordrecht.

Vossler, C. A., Doyon, M., & Rondeau, D. (2012). Truth in consequentiality: theory and field evidence on discrete choice experiments. *American Economic Journal: Microeconomics*, *4*(4), 145-71.

Vossler, C. A., & Watson, S. B. (2013). Understanding the consequences of consequentiality: Testing the validity of stated preferences in the field. *Journal of Economic Behavior & Organization*, *86*, 137-147.

Zawojska, E., Bartczak, A., & Czajkowski, M. (2019). Disentangling the effects of policy and payment consequentiality and risk attitudes on stated preferences. *Journal of Environmental Economics and Management*, *93*, 63-84.