

Substitution Effects and Spatial Preference Heterogeneity in Single- and Multiple-Site Choice Experiments

Ivana Logar Senior researcher, Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

Roy Brouwer Professor, Department of Economics, executive director, The Water Institute, University of Waterloo, Ontario, Canada, and Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

ABSTRACT A novel discrete choice experiment (DCE) design allows testing of substitution effects based on the framing of, and distances to, substitute sites. The same ecosystem services and biodiversity provided at different sites are valued independently in single-site DCEs and simultaneously in a multiple-site DCE, using an identical experimental design. Site-specific utility functions are estimated with and without controlling for spatial preference heterogeneity. Framing choices as alternative projects at single or multiple sites significantly influences substitutability between sites, choice behavior, and welfare estimates. Distance decay displays significant heterogeneity in different directions from the valued sites and between urban and rural areas. (JEL Q51)

1. Introduction

Spatial dimensions have long been ignored in research on the economic valuation of environmental goods and services using stated preference (SP) methods. Recently, however, the role that various spatial factors play in determining public preferences and willingness-to-pay (WTP) values has received increasing attention. Empirical evidence shows that respondents' preferences and welfare measures are not uniformly distributed across space, but display significant spatial variation (Schaafsma, Brouwer, and Rose 2012;

Schaafsma et al. 2013; Jørgensen et al. 2013). Existing studies demonstrate that distance from a respondent's home to both the site being valued and its substitute sites is a relevant source of preference heterogeneity (Jørgensen et al. 2013; Schaafsma et al. 2013). This has important implications for the aggregation and transferability of economic value estimates, since the omission of these spatial characteristics can lead to biased results. Moreover, SP surveys implicitly assume that respondents consider the relevant range of available public and private substitute and complementary goods when valuing the good under study. A number of studies, however, show that asking respondents to value a good independently or next to other relevant goods produces significantly different welfare estimates, highlighting the importance of adequately framing substitutes (Hoehn and Loomis 1993; Cummings, Ganderton, and McGuckin 1994; Brown and Duffield 1995). Most research on spatial preferences and substitution effects has been carried out using contingent valuation surveys (e.g., Neill 1995; Hailu, Adamowicz, and Boxall 2000; Hanley, Schläpfer, and Spurgeon 2003; Bateman et al. 2006). Empirical evidence from discrete choice experiments (DCE) is still rare.

The novelty of this paper is that it captures framing effects resulting from a distinct presentation of substitutes directly in the DCE design. Using an identical experimental design, the same ecosystem services and biodiversity provided at two different sites are valued (1) independently in two separate single-site DCEs (SDCEs), each focusing exclusively on one of the two study sites; and (2) as substitutes in a multiple-site DCE (MDCE),

Land Economics • May 2018 • 94 (2): 302–322
ISSN 0023-7639; E-ISSN 1543-8325
© 2018 by the Board of Regents of the University of Wisconsin System

where each study site is presented as a labeled choice alternative. The MDCE thereby explicitly requires respondents to consider substitutability between the ecosystem services and biodiversity provided at the two sites in their choice process, while in the SDCE these effects are captured through simple reminders about possible substitute sites.

This is, to our knowledge, the first study that examines distance-decay and substitution effects by comparing SDCEs and MDCE. In addition to distances to the valued and to the substitute sites, we control for directional heterogeneity, that is, for the possibility of respondents' preferences not being the same in all directions from the valued site, as proposed by Schaafsma et al. (2013). Such directional effects may result from an uneven spatial distribution of substitute sites. Adding to these previous studies, we also test for differences in distance decay between urban and rural areas. The main objectives of this paper are to investigate (1) to what extent the framing of the river restoration projects in the DCE as choice alternatives for single or multiple sites influences substitution effects, (2) whether SPs and welfare estimates for the same services at the same river site derived from SDCEs and an MDCE are significantly different, and (3) whether welfare functions for identical ecosystem services and biodiversity are transferable between different river sites when controlling for spatial preference heterogeneity.

2. Spatial Preferences, Substitution, and Framing Effects in SP Studies

Bateman et al. (2006), Schaafsma et al. (2013), and Jørgensen et al. (2013) provide comprehensive literature reviews on spatial preferences. By far the most commonly studied spatial phenomenon in the SP literature is distance decay, which refers to the negative relationship between public WTP and the distance between respondents' place of residence and the study site. The farther away respondents live from the study area, the less (likely it is that) they are willing to pay for its im-

provement or preservation (Pate and Loomis 1997). There are various reasons for distance decay (Rolfe and Windle 2012). For example, the actual use of an environmental resource is likely to be lower for people who live farther away from it. The distance-decay effect is, therefore, expected primarily among resource users, while the implications for nonusers are less straightforward (e.g., Bateman et al. 2006). More recently, it has been acknowledged that the assumption of homogenous and continuous distance decay is limiting and potentially erroneous. As a result, an increasing number of SP studies have attempted to allow for more complex spatial patterns in the distribution of stated WTP values (Campbell, Hutchinson, and Scarpa 2009; Schaafsma, Brouwer, and Rose 2012; Schaafsma et al. 2013; Jørgensen et al. 2013; Johnston and Ramachandran 2014). These studies are, however, still exceptions in the SP literature.

Besides being a proxy for the price of environmental nonmarket goods and services, distance decay is closely related to substitution effects, as the set of available substitutes is expected to be larger for respondents who live farther away from the valued resource (Loomis 2000). The opportunity costs of visiting the site under study rise with increasing distance and travel time, resulting in lower demand. At the same time, the opportunity costs of visiting substitute sites become lower and lead to higher demand. Therefore, as the number of available substitutes increases, public WTP for the site of interest is expected to decline at greater distances (Jørgensen et al. 2013). Despite recognizing the importance of substitute goods in shaping spatial preferences, only a handful of SP studies account for the availability of, and distance to, substitutes when analyzing distance effects (Schaafsma, Brouwer, and Rose 2012). Most of these studies find that apart from the distance to the study site, information about substitute sites has a significant effect on WTP for the site being valued (Brown and Duffield 1995; Pate and Loomis 1997; Brouwer and Slangen 1998; Hanley, Schläpfer, and Spurgeon 2003; Concu 2007; Jørgensen et al. 2013; Lizin et al. 2016).

The advantage of using DCEs for investigating framing and substitution effects is that they allow a range of substitute goods to

be included in a choice set simultaneously, forcing respondents to consider these when making trade-offs (Rolfe, Bennett, and Louviere 2002). This is an important feature given that, for example, Neill (1995) found that reminders about substitutes do not affect the economic value of a specific environmental good, contrary to asking respondents to value a good alongside other substitute goods. Studies by Rolfe, Bennett, and Louviere (2002), Schaafsma and Brouwer (2013), Schaafsma, Brouwer, and Rose (2012), Schaafsma et al. (2013), and Lizin et al. (2016) are among the very few that analyze substitution effects by including substitute sites in the DCE design itself. Rolfe, Bennett, and Louviere (2002) have tested framing effects associated with substitute goods in MDCEs, where each location served as a labeled alternative. Their results show that even minor changes in the presentation of substitutes can lead to significant changes in model parameters, which suggests that framing effects related to substitutes may be more widespread in SP studies than is commonly thought. However, Schaafsma and Brouwer (2013) find no significant framing effects in a labeled MDCE when changing the number of substitute sites included in a choice set. Schaafsma, Brouwer, and Rose (2012) and Schaafsma et al. (2013) ask respondents to choose directly between multiple substitute sites, and demonstrate that distance decay is neither homogeneous nor linear across the study area, displaying significant directional effects as a result of differences in the availability of substitutes across the study area. These directional effects are rarely accounted for in SP studies. Lizin et al. (2016) also apply an MDCE to investigate distance-decay and substitution effects simultaneously, but do not consider directional differences in respondents' spatial preferences. Their results show that it is primarily the distance to the substitute site that affects choice behavior, not the distance to the site being valued. Most of the existing SP studies hence focus on unidirectional distance decay in MDCEs and do not allow for directional heterogeneity in their analysis.

Contrary to these previous DCE studies, we apply SDCE in addition to MDCE for identical sites. SDCEs serve as a control group and

allow us to formally test the impact of valuing substitute sites simultaneously compared with the inclusion of simple reminders about the substitutes. Moreover, we capture spatial preference heterogeneity through the distance-decay function that controls for distances to the sites under valuation and their substitute sites, possible directional effects, and differences between urban and rural areas.

3. Case Study, Survey Design, and Sampling Procedure

Case Study Description

Stretches of the rivers Thur and Töss, located in northeastern Switzerland, serve as sites for our case studies. As many other Swiss rivers, the Thur and Töss, both tributaries of the river Rhine, have been channelized over the past centuries, which has led to the degradation of their riverine ecosystems (Woolsey et al. 2005). Nowadays an increasing number of river restoration projects are being undertaken in Switzerland in order to return rivers to their more natural conditions. This is expected to improve their ecological state, prevent further biodiversity loss, and restore lost ecosystem services (Palmer, Menningen, and Bernhardt 2010; Bernhardt et al. 2005). In the revised Water Protection Act, the Swiss government set the goal to restore 4,000 km of rivers over the next 80 years (FOEN 2011; Kurth and Schirmer 2014).

In recent years, specific sections of the rivers Thur and Töss have been restored over a length of 1.5 km and 0.2 km, respectively. The ecological effects of the restoration measures along these stretches are assessed and reported by Paillex et al. (2017). Their findings show that restoration projects along both rivers have led to improvements in landscape and biodiversity, but not in river water quality. The restored river stretches were, furthermore, too short to have an impact on flood risk. Woolsey et al. (2007) report positive effects of restoration on the number of visitors and the range of recreational activities undertaken at the river Thur.

This study elicits local residents' WTP for further restoration measures of specific de-

Table 1
Attributes and Attribute Levels Used in the Discrete Choice Experiment

Attribute	Attribute Levels
Length of the river section that would be restored	0 km, ^a 0.5 km, 1 km, 2 km
Walking along the river	Yes, No ^a
Swimming in the river	Yes, No ^a
Barbecuing on the river bank	Yes, No ^a
Biodiversity	Low (60% of potential plant and animal species present) ^a Medium (75% of potential plant and animal species present) High (100% of potential plant and animal species present)
Price (increase in canton taxes per person per year)	0 CHF, ^a 25 CHF, 50 CHF, 100 CHF, 150 CHF, 250 CHF, 500 CHF

^a Attribute level in the status quo alternative. The status quo levels for the length and price attributes could not appear in the hypothetical choice alternatives. CHF, Swiss francs.

graded sections of the rivers Thur and Töss, located approximately 500 m upstream from the restored river stretches. These degraded sections are representative of the restored river stretches before they were restored (Paillex et al. 2017). According to expert judgment, the effects of restoring these degraded river sites are expected to be the same as at the restored sites. The incremental effects of upstream restoration on the restored river stretches downstream would be negligible, since both the restored sections and the sections that might be restored are rather short and disconnected. These features, combined with the hydraulic dynamics of the rivers, imply that the new restoration projects would not result in notable impacts downstream. The lack of connectivity between the restored river stretches and the stretches that might be restored means, for example, that biodiversity would evolve independently at each site. Although some differences exist between the Thur and Töss river sites, they provide identical ecosystem services, which are being valued in the DCE. The river Thur is somewhat wider and longer than the Töss. The Thur study site is surrounded mainly by agricultural land with certain parts covered by alluvial forests, while the site to be restored at the river Töss is located within a forested area. The distance between the two river sites is 15 km. A map of the study area is presented in [Appendix Figure A1](#). The map shows the location of the two study sites, surrounding land use, population density, and other rivers in the area, which serves for illustrating the spatial distribution of substitute sites and the overall river restoration potential in the region.

Survey and Experimental Design

The attributes and attribute levels employed in the DCEs are presented in Table 1. They were selected in close collaboration with the same experts who carried out the environmental impact assessment of the past restoration projects along the two rivers (Paillex et al. 2017).

The first choice attribute defines the length of the river section that would be restored and is measured in kilometers. The next three choice attributes capture the effect of river restoration on the most common recreational activities among river users, namely, walking, swimming, and barbecuing. The levels of these three attributes are defined as either having or not having the possibility to undertake the activity. The fifth attribute, biodiversity, reflects the expected improvement in the biological state of the river that would result from further restoration measures. The attribute levels are defined in terms of the number of plant and animal species that can be found in and around the river. Low, medium, and high biodiversity levels indicate to what extent the present number of species compares with their maximum potential number. The current biodiversity level is low for both rivers, meaning that only 60% of all potential species are present. This can be increased to 75% (medium level) or 100% (high level). The last attribute is a tax increase, with levels ranging from 25 to 500 Swiss francs (CHF) per person per year. An increase in the annual canton taxes came out as the most credible payment vehicle after pretesting because most taxes in Switzerland are paid once per year and river

restoration projects fall under the jurisdiction of the cantons.

A D-efficient DCE design was generated in the software Ngene, using priors from the pretest data (e.g., Rose et al. 2008). This resulted in 36 different choice tasks, which were blocked into six choice sets, consisting of six choice tasks each. The choice sets were randomly distributed across respondents. Choice tasks were presented on a one-page choice card. Each choice card included two policy alternatives and the status quo alternative. This latter “opt-out” meant that no further river restoration measures would take place, and that canton taxes would not increase. Respondents were asked to choose their most preferred alternative in each choice task. If they consistently chose the opt-out in all six choice tasks, a follow-up question asking about their underlying reasons enabled filtering out protest responses (Brouwer and Martín-Ortega 2012).

To test our hypotheses, specified in Section 4, three different versions of DCEs were used and administered to three random samples of respondents. Two SDCE versions were applied to the rivers Thur and Töss separately. Each SDCE version, hence, focused exclusively on either the Thur or the Töss river site. Apart from targeting different study sites, the two SDCE versions were identical. A map of the study area similar to that in [Appendix Figure A1](#) was presented to respondents before the choice tasks. The map indicated exactly which section of the river Thur or Töss would be restored, and the accompanying text informed the respondents that the choice alternatives describe different impacts of the restoration projects at that particular river site. For this reason, there was no need to specify the river name in the choice alternatives in the two SDCE versions, and these were instead simply labeled as options A and B. Therefore, although the choice alternatives were unlabeled, they implicitly included the river labels, as it was explained to respondents that they refer to the restoration of a specific section of the river Thur or the river Töss. In the two SDCEs, respondents were, hence, asked to choose between alternative river restoration projects at a specific single river site.

The third DCE version included the MDCE, in which respondents were asked to

choose between identical restoration projects for the same Thur and Töss river sites as in the SDCEs, but here each river site was represented as a labeled alternative. In the MDCE, the first hypothetical alternative was always labeled as the restoration of the Thur and the second hypothetical alternative as the restoration of the Töss. The third alternative was labeled as “no change” and represented the status quo situation where no further restoration measures would be implemented for either of the two rivers. All three DCE versions (i.e., the two SDCEs and the MDCE) included the same reminder of other substitute sites and the respondent’s household budget constraint, as recommended by NOAA (1993). The respondents in all three samples were also informed that only one river restoration project would eventually be implemented, thereby excluding the possibility of having multiple restoration projects or restoring both the Thur and the Töss sites. An example of the choice card used in the SDCEs and MDCE is presented in Figure 1.¹

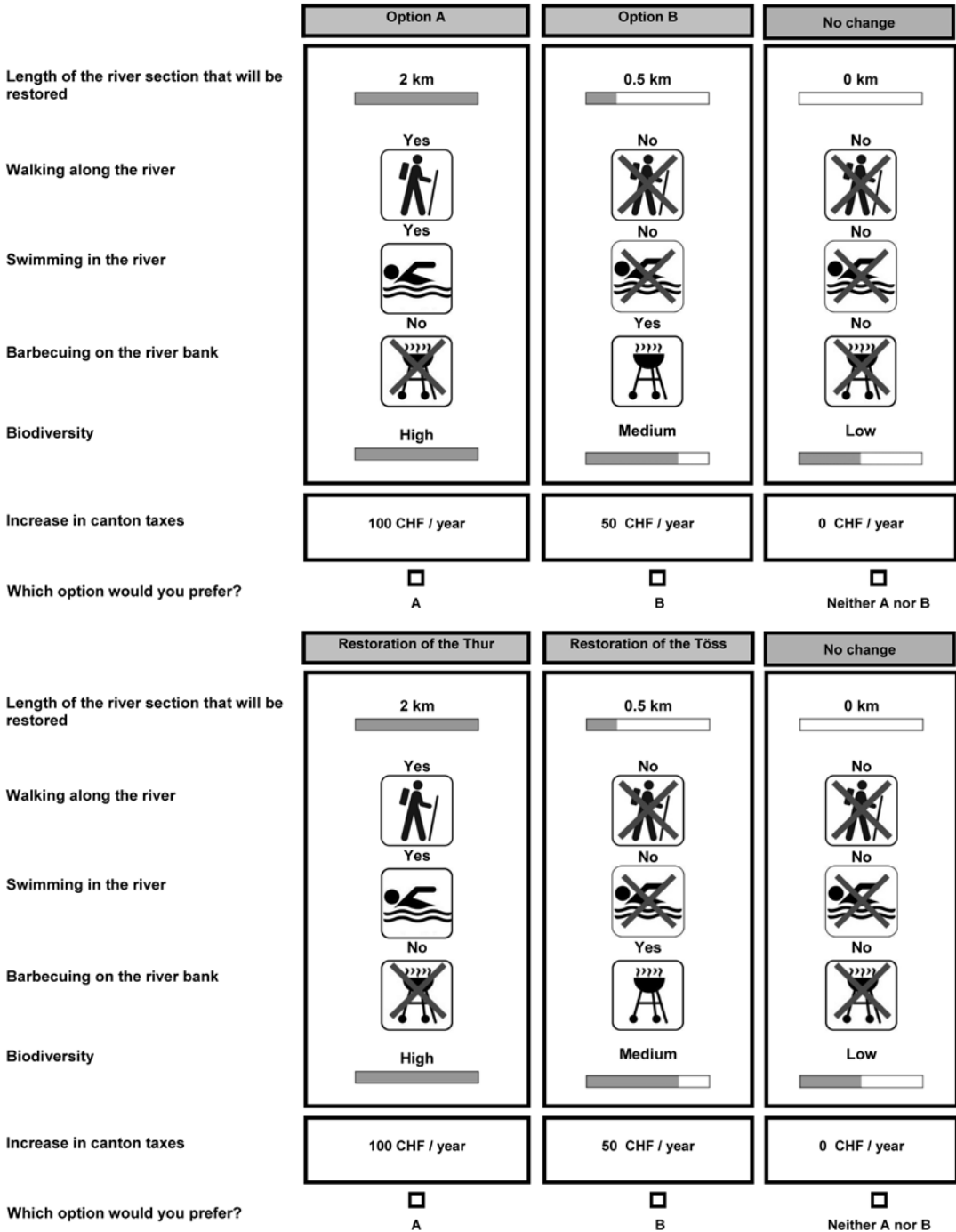
The survey was thoroughly pretested in the study area through in-person interviews. The final survey was conducted in March 2015 and administered in person in public areas by a professional marketing agency.

Sampling Procedure

A spatial sampling strategy is of utmost importance for investigating spatial preference heterogeneity. Detecting how preferences vary across space requires a sample that represents the spatial distribution of the population (Concu 2007). The target population in this study consists of households located within a 35 km radius from (1) the river section that might be restored in the future in the SDCEs and (2) the midpoint between the two sites in the MDCE. This sampling strategy resulted in three different samples, which partly overlap in space. The complete sample includes respondents from 114 towns and villages in the cantons of Zurich and Thurgau, located at varying distances from the Thur and

¹Note that the choice alternatives refer to the restoration of the river Thur in one SDCE sample and to the restoration of the river Töss in the other SDCE sample.

Figure 1
 Choice Card Example for the Two Single-Site Discrete Choice Experiments (*above*)
 and the Multiple-Site Discrete Choice Experiment (*below*)



Töss study sites. The sampling procedure ensured that the three samples are representative of the population in the study area in terms of gender, age, and population density. The latter implied some degree of oversampling of respondents in larger towns and the cities of Zurich and Winterthur. The resulting spatial distribution of respondents for each sample is presented in [Appendix Figure A2](#). Each subsample comprises 250 respondents, so in total 750 individuals were surveyed.

4. Econometric Model and Hypothesis Testing

Choice Models

Choice modeling is based on Lancaster’s (1966) multiattribute utility theory, which assumes that consumers’ utilities for goods can be decomposed into utilities for their characteristics. Preferences for the policy alternatives in the DCE are modeled in terms of McFadden’s (1974) random utility model. The level of utility (U_{ijt}) that an individual i obtains from the alternative j in a choice situation t can be decomposed into a deterministic part (V_{ijt}) and a stochastic part (ε_{ijt}):

$$U_{ijt} = V_{ijt} + \varepsilon_{ijt}. \tag{1}$$

The deterministic component of alternative j can be specified as a linear function of its characteristics or attributes (x_{ijt}) and other explanatory variables (z_i) for every alternative in the choice set $j = (1, \dots, J)$. The utility function takes the following form, where β and γ are the vectors of parameters associated with \mathbf{x} and \mathbf{z} , respectively:

$$U_{ijt} = \beta_i' x_{ijt} + \gamma' z_i + \varepsilon_{ijt}. \tag{2}$$

The probability P that an individual i chooses an alternative j over alternative k from a choice set C is

$$P_{it}(j|C) = P(U_{ijt} \geq U_{ikt}, \forall k \in C, j \neq k). \tag{3}$$

In order to estimate equation [3], we apply a mixed logit model, which is more flexible and less restrictive than a standard conditional

logit model (Train 2003). By enabling coefficients to vary randomly across individuals, the mixed logit model is able to account for preference heterogeneity between respondents. Hence, the model captures unobserved heterogeneity in the choice attributes but may fail to explain the sources of heterogeneity (Hynes, Hanley, and Scarpa 2008). One way to overcome this limitation is to include interactions of choice attributes with respondent-specific characteristics in the utility function (Revelt and Train 1998). These interactions can detect sources of heterogeneity, while controlling for unobserved heterogeneity, and at the same time improve the model fit.

In the mixed logit model, the probability of observing a sequence of choices J_i for an individual i is conditional on the parameter vectors β_i and γ . The scaling factor λ denotes the scale of the utility. The likelihood function is specified as the product of conditional probabilities over all choice tasks $t = (1, \dots, T)$ presented to the respondent:

$$P(J_i | \beta_i, \gamma) = \prod_{t=1}^T \left\{ \frac{\exp[\lambda(\beta_i' x_{ijt} + \gamma' z_i)]}{\sum_{k \in C} \exp[\lambda(\beta_i' x_{ikt} + \gamma' z_i)]} \right\}. \tag{4}$$

Since the researcher does not observe β_i , equation [4] is integrated over all the possible values of β_i using the density function. The resulting unconditional choice probability is

$$P(J_i | \theta, \gamma) = \int P(y_i | \beta_i, \gamma) f(\beta_i | \theta) d\beta_i. \tag{5}$$

The vector of parameters is estimated by maximizing the likelihood function. The integral in equation [5] generally does not have a closed form, and thus it cannot be solved analytically and requires simulation. The probability is approximated through a maximum simulated likelihood, which generates draws from distributions with given means and standard deviations. Halton sequences have been found to produce more precise results than independent random draws in the estimation of mixed logit models (Bhat 2001).

Spatial Preference Heterogeneity

Cameron (2006) demonstrates that controlling only for a simple unidimensional distance in-

indicator, which implicitly assumes that the distance-decay effect is the same in all directions from the site, and ignoring possible directional effects, might lead to insignificant or biased distance parameters. This finding is confirmed by Schaafsma et al. (2013), who build on this earlier work and develop a procedure to account for directional heterogeneity in DCEs. We follow their approach and further expand the spatial analysis by distinguishing between distance-decay effects in urban and rural areas. The distance-decay function for the specific site with coordinates (x,y) is specified as follows:

$$f(x,y) = \beta_n \text{Distance}_{in} + \delta_n (\text{Distance}_{in} \times \text{Urban}_i) + (\varphi_{1n} \text{Longitude}_{in} + \varphi_{2n} \text{Latitude}_{in}) + (\eta_{1n} \cos \theta_{in} + \eta_{2n} \sin \theta_{in}), \quad [6]$$

where Distance_{in} denotes the one-way road distance from respondent i to the valued river site n . $\text{Distance}_{in} \times \text{Urban}_i$ is the interaction term between road distance and a dummy variable indicating whether the respondent lives in an urban or rural area. Longitude_{in} and Latitude_{in} are the longitudinal and latitudinal Euclidean distances between the site under valuation and the respondents' place of residence, which capture differences in distance decay between east and west and north and south, respectively. Cosine (sine) uses the angle θ_{in} between the study site and the respondent's place of residence, measured in degrees counterclockwise starting from the east, in order to reflect differences in preferences of respondents who live east and west (north and south) from the site, independently of distance.

Main Hypotheses

The main objective of this study is elaborated in four hypotheses. The first hypothesis tests whether any framing effect can be detected due to the inclusion of two substitute sites directly into the choice set in the form of labeled alternatives. To this end, we specify a null hypothesis, which states that marginal WTP ($MWTP$) values derived from the SDCEs are equal to those estimated from the MDCE for

the same ecosystem service and biodiversity provision levels at the same river site:

$$H_0^{1a} : MWTP_{\text{Thur}}^{\text{SDCE}} = MWTP_{\text{Thur}}^{\text{MDCE}}, \quad [7]$$

$$H_0^{1b} : MWTP_{\text{Töss}}^{\text{SDCE}} = MWTP_{\text{Töss}}^{\text{MDCE}}. \quad [8]$$

Poe, Giraud, and Loomis's (2005) test procedure will be applied to test this first hypothesis. In theory, for substitute goods, marginal utility decreases with additional units of similar goods, implying that the welfare measure for a good valued on its own is likely to be greater than WTP for the same good when valued in a bundle of other goods (Mas-Colell, Whinston, and Green 1995; Varian 2014). On the contrary, if the two goods are complements, their combined benefits, when valued jointly, exceed the sum of their individually and independently assessed values (Hoehn and Loomis 1993; Brown and Duffield 1995; Hailu, Adamowicz, and Boxall 2000). Note, however, that in our experimental setting the two goods are not valued jointly, since we do not elicit a single WTP estimate for both river sites but ask respondents to value them next to each other. Therefore, we cannot test this formally in an alternative hypothesis.

The second hypothesis tests the existence of distance-decay effects. Here, a distinction is made between conventional unidirectional and spatially heterogeneous multidirectional distance decay. The unidirectional distance decay is tested using the estimated parameter associated with the distance at which respondents live from the river site under valuation, which is expected to be significantly different from zero and have a negative sign. The corresponding alternative hypothesis is

$$H_a^{2a} : \beta_{\text{Distance to valued river site}} < 0. \quad [9]$$

Multidirectional distance decay is present if the estimated coefficients associated with the longitudinal and latitudinal distances between the respondents' place of residence and the river sites under valuation are significantly different from zero and negative. Therefore, the following additional hypotheses are tested:

$$H_a^{2b} : \beta_{\text{Longitudinal distance to valued river site}} < 0, \quad [10]$$

$$H_a^{2c} : \beta_{\text{Latitudinal distance to valued river site}} < 0. \quad [11]$$

The third hypothesis tests the presence of substitution (or complementarity) effects between the two river sites by including the distance to the alternative river site in the utility function for the river under valuation. The null hypothesis is in this case specified as follows:

$$H_0^3 : \beta_{\text{Distance to alternative river site}} = 0. \quad [12]$$

Substitution implies that the river sites provide similar services that can replace each other, while the presence of complementarity means that the services provided at the two sites contribute to each other and, when combined, create added value. Formally, substitute (complementary) goods have a positive (negative) cross-elasticity of demand. This means that if the price of one good increases, the demand for its substitute (complement) increases (decreases). The rationale behind this mechanism is that a substitute good can replace the original good, while complementary goods experience joint demand. Analogously, using distance to the alternative river site as a proxy for cross-price effects and WTP for the valued site as an indicator of its own demand, a positive (negative) correlation between the two variables is expected for substitute (complementary) river sites. Therefore, if two river sites are substitutes (complements), the estimated parameter for the distance to the alternative river site is expected to be significantly different from zero and positive (negative) (Schaafsma et al. 2013). That is, the greater the distance to the alternative river site with substitutable (complementary) characteristics, the higher (lower) is the expected value that respondents will assign to the ecosystem services and biodiversity provided by the river site under valuation. If two river sites provide substitute services, the associated alternative hypothesis is specified as follows (Schaafsma et al. 2013):

$$H_{a1}^3 : \beta_{\text{Distance to alternative river site}} > 0. \quad [13]$$

If two river sites provide complementary services, the alternative hypothesis is

$$H_{a2}^3 : \beta_{\text{Distance to alternative river site}} < 0. \quad [14]$$

The second and the third hypotheses are tested based on the outcomes of the estimated SDCE and MDCE choice models.

Although the two rivers provide the same ecosystem services and biodiversity, they are distinct in several aspects (width, length, and surrounding land use). This could lead to differences in respondents' preferences for restoring the two river sites and hinder their transferability. Our last hypothesis is hence specified as an alternative hypothesis:

$$H_a^4 : \beta_{\text{Thur}} \neq \beta_{\text{Töss}} \quad [15]$$

If the hypothesis cannot be rejected, the estimated preference parameters β in the utility functions for the two rivers are significantly different. To test this hypothesis between the two SDCEs, a procedure proposed by Swait and Louviere (1993) is applied. In addition, the preference parameters between the two river-specific utility functions estimated from the MDCE is compared using the Wald test.

5. Results

Sample Characteristics

Statistical tests show that there are no differences at the 1% significance level between the three samples with respect to the most relevant sociodemographic characteristics for explaining stated choices, such as gender, age, income, and membership in environmental organizations. Significant differences are found only for education and household size.

Distances from respondents' homes to the two rivers and other spatial trend variables are calculated in ArcGIS² using their postal codes. The unidirectional distance variables measure the one-way road distances between the respondent's place of residence and the midpoint of the river section that would be restored. The survey also collected information about respondents' use of the rivers under valuation, their knowledge about river restoration projects, and their perception of the effects of previ-

² See ArcGIS, provided by Esri, available at www.arcgis.com/features/index.html.

ous river restoration projects. An overview of the socioeconomic characteristics and other descriptive statistics of the three samples is presented in [Appendix Table A1](#).

The average road distance to the river Thur is higher than to the river Töss in all three samples. The main reason for this is that the Thur is located farther away than the Töss for the residents of Zurich and Winterthur, who represent one-third of the overall sample. A *t*-test shows that the differences in average road distance between the three samples are statistically significant at the 1% significance level. These differences can potentially influence choice behavior and will hence have to be controlled for in the choice models. The most frequent recreational activities that respondents undertake when visiting the rivers under valuation match the selected choice attributes for recreational opportunities that further river restoration would offer.

Estimated Choice Models

Around 4% of the choices in all three samples are classified as protest responses and excluded from the analysis, following common practice in the SP literature (Brouwer and Martín-Ortega 2012). This low protest rate indicates that the trade-offs respondents were asked to consider in the DCE worked well. In the SDCEs, which focus on one particular river site, respondents are not expected to have a specific preference for one of the two alternative river restoration projects. Alternatives A and B are indeed chosen almost equally often within each SDCE sample: each of the two alternatives accounts for 25% and 33% of the choices in the Thur and Töss SDCE version, respectively. In the MDCE, respondents' preferences for one river over the other might lead to one alternative being chosen more frequently than the other. However, we do not find any evidence for this, as restoration of the Thur and Töss is chosen equally often in 36% and 37% of the cases, respectively. Substantial differences in choice behavior are found between the three samples when examining the share of opt-out responses. Opt-out choices are considerably higher in the SDCE sample focusing on the Thur (49%) than in the SDCE sample focusing on the Töss (34%), and are

lowest in the MDCE sample (27%). These differences in opt-out shares are expected to affect the choice parameters across the three samples.

Table 2 describes the covariates included in the choice models. Socioeconomic characteristics of respondents commonly found in SP studies, such as age, household income, or membership in an environmental organization, turned out to play no significant role in explaining SPs for the restoration projects and are, therefore, not included in the final models.

The results of the mixed logit choice models for the three samples are presented in Table 3. These extended models include covariates that account for spatial preference heterogeneity. The attribute-only models serve as a baseline for examining the effects of the covariates on choice parameters, and are included in [Appendix Table A2](#). In the MDCE model, river-specific utility functions are estimated by specifying a separate set of model parameters for each choice alternative. The models are estimated using a maximum simulated likelihood procedure with 500 Halton draws in NLOGIT 5.0.³

The McFadden pseudo- R^2 statistics show a good overall fit of the estimated models. They are similar to, or even outperform, those reported in other spatial DCE studies (e.g., Concu 2007; Rolfe and Windle 2012; Lizin et al. 2016). Alternative-specific constants (ASCs) are included in the utility functions for the two experimentally designed alternatives. ASCs and preference parameters for the choice attributes are specified as random variables in the estimation procedure, which allows them to vary across respondents and capture preference heterogeneity. The coefficients for the choice attributes coded as dummy variables, namely, recreational activities and biodiversity, are specified to be uniformly distributed, as recommended by Hensher, Rose, and Greene (2005). The coefficients for the ASCs, the length of the river section that would be restored, and price are assumed to follow a normal distribution, as this proved to generate the best statistical fit.

³See NLOGIT, provided by Econometric Software, available at www.limdep.com.

Table 2
Explanation of the Variables Used in the Choice Models

Variable	Description	Coding and Units
Length of the river section that would be restored	Choice attribute; indicates the length of the river restoration project	Continuous variable; in kilometers
Walking along the river	Choice attribute; indicates whether river restoration would provide an option to walk along the river	Dummy variable, 1 = Walking along the river is possible
Swimming in the river	Choice attribute; indicates whether river restoration would provide an option to swim in the river	Dummy variable, 1 = Swimming in the river is possible
Barbecuing on the river bank	Choice attribute; indicates whether river restoration would provide an option to barbecue on the river bank	Dummy variable, 1 = Barbecuing on the river bank is possible
High biodiversity	Choice attribute; indicates whether river restoration would increase biodiversity to a high level	Dummy variable, 1 = High biodiversity level is achieved (100% of potential plant and animal species are present)
Medium biodiversity	Choice attribute; indicates whether river restoration would increase biodiversity to a medium level	Dummy variable, 1 = Medium biodiversity level is achieved (75% of potential plant and animal species are present)
Price	Choice attribute; indicates an increase in annual canton taxes per person for the purpose of river restoration	Continuous variable; in Swiss francs
Walking attribute × Respondents who walk when visiting the rivers	Interaction term between the choice attribute for walking and a dummy variable indicating whether the respondent usually walks when visiting the rivers	Dummy variable, 1 = Walking along the river is possible and respondent usually walks when visiting the rivers
Swimming attribute × Respondents who swim when visiting the rivers	Interaction term between the choice attribute for swimming and a dummy variable indicating whether the respondent usually swims when visiting the rivers	Dummy variable, 1 = Swimming in the river is possible and respondent usually swims when visiting the rivers
Barbecuing attribute × Respondents who barbecue when visiting the rivers	Interaction term between the choice attribute for barbecuing and a dummy variable indicating whether the respondent usually barbecues when visiting the rivers	Dummy variable, 1 = Barbecuing on the river bank is possible and respondent usually barbecues when visiting the rivers
ln(Road distance to the Thur + 1)	Natural logarithm of road distance between respondent's home and the river Thur	Continuous variable; road distance in kilometers
Road distance to the Töss	Road distance between respondent's home and the river Töss	Continuous variable; in kilometers
Road distance to the valued river × Respondents living in urban areas	Interaction term between the linear road distance to the valued river and a dummy variable indicating whether the respondent lives in an urban area (towns with >20,000 inhabitants)	Continuous variable; in kilometers
Longitudinal distance to the valued river	Euclidean (air) distance between respondent's home and the valued river site in the East-West direction	Continuous variable; in kilometers; positive (negative) values indicate East (West)
Latitudinal distance to the valued river	Euclidean (air) distance between respondent's home and the valued river site in the North-South direction	Continuous variable; in kilometers; positive (negative) values indicate North (South)
$\cos\theta$	Cosine of the angle between respondent's home and the valued river site	Continuous variable; range: -1 to 1; positive (negative) values indicate East (West)
$\sin\theta$	Sine of the angle between respondent's home and the valued river site	Continuous variable; range: -1 to 1; positive (negative) values indicate North (South)

Preferences for the Choice Attributes

The results for the choice attributes in the extended models remain fairly robust compared

with the attribute-only models. The standard deviations of the random parameters in Table 3 are highly significant for all choice attri-

Table 3
Estimated Mixed Logit Choice Models for the Three Samples

Variable	SDCE	SDCE	MDCE	
	(1) Thur	(2) Töss	(3) Thur	(4) Töss
<i>Mean Estimates of Random Parameters</i>				
Constant	-1.690 (6.010)	-27.741*** (10.103)	-41.346** (16.331)	5.140 (12.584)
Length of the river section that would be restored	1.161** (0.455)	0.617** (0.255)	7.828*** (2.142)	6.905*** (2.114)
Walking along the river	3.471*** (1.110)	2.042*** (0.532)	26.927*** (6.846)	2.411 (1.991)
Swimming in the river	1.112* (0.609)	0.923** (0.387)	9.716*** (2.605)	4.680** (2.100)
Barbecuing on the river bank	1.917*** (0.717)	-0.048 (0.411)	2.520 (1.701)	1.280 (1.365)
High biodiversity	1.936** (0.838)	0.752 (0.493)	15.703*** (4.312)	11.947*** (3.894)
Medium biodiversity	0.979 (0.786)	0.454 (0.508)	13.282*** (3.554)	12.814*** (4.185)
Price	-0.062*** (0.017)	-0.024*** (0.004)	-0.081*** (0.020)	-0.081*** (0.021)
<i>Nonrandom Parameters</i>				
Walking attribute × Respondents who walk when visiting the rivers	3.400*** (1.303)	2.005*** (0.573)	2.078 (2.165)	16.740*** (4.576)
Swimming attribute × Respondents who swim when visiting the rivers	2.923*** (0.880)	1.791*** (0.528)	7.780*** (2.558)	6.396** (2.602)
Barbecuing attribute × Respondents who barbecue when visiting the rivers	3.985*** (1.046)	2.101*** (0.617)	18.774*** (4.835)	7.586** (3.200)
ln(Road distance to the Thur + 1)	-7.572** (3.766)	10.393*** (3.468)	4.404 (6.091)	-3.248 (4.578)
Road distance to the Töss	0.887*** (0.276)	-0.465*** (0.101)	-0.865*** (0.289)	-1.033*** (0.282)
Road distance to the valued river × Respondents living in urban areas	0.871*** (0.240)	0.173*** (0.058)	5.511*** (1.595)	0.588*** (0.198)
Longitudinal distance to the valued river	-0.101 (0.070)	-0.109 (0.077)	-0.805*** (0.222)	-1.053*** (0.295)
Latitudinal distance to the valued river	-0.829*** (0.270)	0.106 (0.116)	0.208 (0.251)	-0.871*** (0.289)
cosθ	0.571 (0.714)	-0.792 (0.652)	-7.696*** (2.180)	0.315 (0.885)
sinθ	1.346 (0.845)	0.562 (0.486)	2.916*** (1.203)	3.077*** (1.118)
<i>Standard Deviations of Random Parameters</i>				
Constant	7.090*** (1.727)	4.226*** (0.736)	6.601*** (1.850)	7.266*** (1.844)
Length of the river section that would be restored	2.340*** (0.580)	1.389*** (0.310)	10.456*** (2.663)	10.684*** (2.571)
Walking along the river	8.199*** (2.039)	4.719*** (0.804)	39.451*** (9.658)	27.737*** (7.135)
Swimming in the river	4.049*** (1.064)	3.694*** (0.713)	31.070*** (7.867)	9.686*** (2.441)
Barbecuing on the river bank	4.973*** (1.448)	4.863*** (0.774)	34.740*** (8.600)	28.724*** (7.532)
High biodiversity	6.163*** (1.617)	2.261*** (0.751)	22.656*** (5.363)	3.357** (1.641)

(table continued on following page)

Table 3 (continued)

Variable	SDCE	SDCE	MDCE	
	(1) Thur	(2) Töss	(3) Thur	(4) Töss
Medium biodiversity	2.992** (1.456)	0.726 (0.676)	18.480*** (4.695)	0.091 (2.206)
Price	0.055*** (0.016)	0.026*** (0.004)	0.076*** (0.019)	0.102*** (0.026)
Log-likelihood	-781.67	-893.08		-1,004.53
Likelihood ratio test (χ^2)	1,508.37***	1,311.94***		1,154.94***
Pseudo R^2	0.491	0.423		0.365
AIC/N ^a	1.155	1.304		1.467
Number of observations	1,398	1,410		1,440

Note: Standard errors are reported in parentheses. MDCE, multiple-site discrete choice experiment; SDCE, single-site discrete choice experiment.

^a Akaike information criterion with a correction for sample size.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

butes, except for medium biodiversity at the river Töss in the SDCE and MDCE models, revealing that preferences for the choice attributes differ across individual respondents.

The significant negative coefficient estimates for the ASCs in the SDCE model for Töss and in the MDCE model for Thur suggest that without consideration of the choice attributes, the hypothetical alternatives decrease the utility of respondents compared with the status quo alternative. Respondents who received the SDCE version for the Thur and those who prefer restoration of the river Töss in the MDCE seem indifferent between the two hypothetical alternatives and the status quo. These findings correspond with the observed relatively large number of opt-out choices in the samples. There exists, however, significant variation across individual respondents in the value they attach to the hypothetical alternatives, as is evident from the significant standard deviation of the ASCs in all four utility functions. The significant positive coefficients for river length imply that respondents value river restoration more if the river stretch to be restored is longer. The coefficients for the price attribute are, as expected, negative and highly significant in all three choice models. They indicate that the higher the increase in canton taxes, the lower the utility of respondents.

For the choice attributes that capture the option to undertake on-site recreational activi-

ties, both main and interaction effects are estimated. The interaction terms between each of the recreational activities and a dummy variable indicating whether respondents undertake this particular activity when visiting the rivers seem to drive the main changes found in the coefficient estimates for the recreational choice attributes between attribute-only and extended models. Accounting for differences between river users and nonusers renders some of the coefficient estimates in the extended models insignificant compared to the attribute-only models. The main effects indicate that respondents are more likely to choose river restoration projects that facilitate walking along and swimming in the river, as reflected by the significant positive coefficients for these recreational attributes, except for the insignificant coefficient estimate for walking along the Töss in the MDCE. Therefore, in general, respondents value restoration measures more if they provide additional recreational opportunities. As expected, users of specific recreational activities value their provision significantly higher than nonusers. Barbecuing is valued significantly only at the Thur in the SDCE if respondents also engage in this activity.

The most noticeable difference between the SDCEs and the MDCE is found in respondents' preferences for improvements in biodiversity levels. These changes in biodiversity levels capture both use and nonuse values.

Table 4
Extended Models: Marginal Willingness-to-Pay Estimates per Person per Year, in Swiss Francs

Choice Attribute	SDCE Thur	SDCE Töss	MDCE		Between SDCE and MDCE Thur	Between SDCE and MDCE Töss
			Thur	Töss		
Length of the river section that would be restored (in kilometers)	18.74 (7.37)	26.19 (12.42)	96.09 (21.03)	85.74 (19.21)	$p < 0.01$	$p < 0.01$
Walking along the river	56.05 (18.52)	86.62 (24.89)	330.54 (44.44)	29.94 (25.54)	$p < 0.01$	$p < 0.05$
Swimming in the river	17.96 (9.79)	39.17 (18.04)	119.27 (28.11)	58.12 (24.29)	$p < 0.01$	$p = 0.26$
Barbecuing on the river bank	30.96 (11.50)	-2.05 (17.33)	30.93 (22.32)	15.90 (17.35)	$p = 0.48$	$p = 0.23$
High biodiversity	31.27 (12.59)	31.91 (23.43)	192.76 (36.70)	148.35 (38.53)	$p < 0.01$	$p < 0.01$
Medium biodiversity	15.81 (12.63)	19.25 (22.91)	163.05 (29.88)	159.12 (41.82)	$p < 0.01$	$p < 0.01$

Note: Standard errors are reported in parentheses. MDCE, multiple-site discrete choice experiment; SDCE, single-site discrete choice experiment.

The MDCE sample exhibits significant preferences for enhancing biodiversity to medium and high levels, while choice behavior in the SDCE samples appears generally unaffected by improvements in biodiversity, with the exception of reaching high biodiversity at the river Thur. It is not possible to attribute these differences to the DCE framing in view of the observed differences between samples and the somewhat lower share of river users in the MDCE sample compared with the SDCE samples (see the discussion of sample characteristics earlier in this section).

Comparison of Welfare Estimates

The MWTP values for the choice attributes in the three samples are presented in Table 4. The estimated MWTP values based on the attribute-only models are included in [Appendix Table A3](#) and again serve as a benchmark for the extended models. Associated standard errors are calculated using the delta method (Greene 2002).

An important first observation is that the welfare estimates derived from the SDCEs are consistently lower than those obtained from the MDCE. This is most likely the result of the higher opt-out shares in the two SDCE samples. Exceptions are the MWTP for walking along the river Töss and for barbecuing at both river sites. Corresponding to the estimated preference parameters in SDCE models in Table 3, the welfare estimates for

biodiversity improvements obtained from the two SDCE versions are not significant, except for achieving a high biodiversity level at the river Thur. The welfare estimates associated with high and medium biodiversity levels are statistically significant for both study sites based on the MDCE model.

Poe, Giraud, and Loomis's (2005) test is used to verify whether the observed differences between the SDCE and MDCE welfare measures are statistically significant for the same river site. The test results are presented in the last two columns in Table 4. MWTP values derived from the MDCE are either significantly higher than or equal to those elicited in the SDCE for the same site. Hence, the first hypothesis on the equality of welfare estimates is rejected for most of the choice attributes for both river sites, implying that a significant framing effect exists between the SDCEs and the MDCE.

An important second observation is that when examining MWTP differences between the two rivers, the MWTP estimates are generally significantly lower for the river Töss than for the Thur in the extended models, while the reverse is observed based on the attribute-only models. In the extended SDCE models, a significant difference between the two river sites is found in the welfare estimates for barbecuing and in the MDCE model for walking and swimming. The MWTP values for barbecuing and swimming differ, however, only at

the 10% significance level. Fewer differences between the sites are found based on the extended models that control for spatial preference heterogeneity.

Spatial Preferences

Each extended choice model includes two unidirectional distance variables, which measure the road distances from the respondents' place of residence to the two river sites. Depending on the specific utility function in which they are included, for the river Thur or river Töss, these variables represent the distance to either the river site under valuation or the alternative river site. In the utility functions for the river Thur based on the SDCE and MDCE versions (columns 1 and 3 in Table 3, respectively), the distance to the Thur captures distance-decay, whereas the distance to the Töss aims to detect the presence of a substitution or complementarity effect between the two river sites. In the SDCE and MDCE utility functions for the river Töss (columns 2 and 4 in Table 3, respectively), these variables play the opposite roles. A significant conventional unidirectional distance-decay effect is found in the SDCE model for the river Thur and in both the SDCE and the MDCE choice models for the river Töss.

A number of distance variable transformations were tested to assess spatial preferences. Those most commonly found in the literature include linear, logarithmic, quadratic, square root, and reciprocal functional forms (e.g., Cameron 2006; Concu 2007). The natural logarithmic specification of the distance to the Thur and the linear form for the distance to the Töss resulted in the best statistical model fit. The logarithmic functional form of distance to the Thur in the utility functions for the Thur captures a distance-decay pattern in which WTP values decline quickly at shorter distances from the study site, and then level out at greater distances. The same variable in the utility functions for the Töss suggests that the substitution or complementarity effect is stronger among respondents who live closer to the Thur and fades away at greater distances.

Coefficient estimates for the interaction terms between the distances to the river sites under valuation and the dummy variable for urban areas are, furthermore, highly

significant across all choice models, meaning that distance-decay functions differ significantly between urban and rural populations.⁴ Their positive signs indicate that at increasing distances from the river sites under valuation, respondents living in urban areas attach a significantly higher value to river restoration than respondents residing in rural areas. As a result, WTP values for river restoration decline at a considerably slower rate among respondents who live in urban areas. This finding is in line with theoretical expectations and can be explained by the relative scarcity of substitute sites in urban compared to rural areas. A similar result is reported by Johnston and Ramachandran (2014), who find that areas with a higher population density are associated with a higher WTP for river restoration. In our study, the result is likely driven by respondents residing in the city of Zurich, which is located farther away from the two river sites than most of the rural areas in this study.

The extended choice models indicate that the estimated distance-decay effects and WTP values are not spatially homogeneous, but instead display significant variation across different directions from the study site. Directional effects are somewhat more prominent in the MDCE than in the SDCEs. Except for the SDCE model for the river Töss, significant multidirectional distance-decay effects exist, either in an east-west or a north-south direction, or both. This is observable from the significant negative coefficients for the longitudinal and latitudinal distances to the river sites under valuation. The parameter estimates for the longitudinal distances are significant in the utility functions for both rivers in the MDCE sample. This means that the farther east of the river sites respondents live, the less likely they are to choose any of the river restoration alternatives. The latitudinal distance variable, which captures distance-decay heterogeneity in the north-south direction, is significantly negative in the SDCE choice model for the

⁴No significant differences in distance-decay effects could be detected between users and nonusers, possibly because current nonusers might become future users after restoration.

river Thur and in the MDCE model for the Töss. This indicates that the distance-decay effect is stronger for respondents living north of the river site to be restored than for those living south of the site. The second hypothesis can, hence, not be rejected in view of the fact that either a unidirectional or multidirectional distance-decay effect is found in all four utility functions.

Further examination of Table 3 reveals that the cosine intercept shifter is significant and negative in the MDCE model for the river Thur. This means that respondents who live west of the Thur, where $\cos(180^\circ) = -1$, are willing to pay significantly more for its restoration compared with respondents living east of this river, where $\cos(0^\circ) = 1$. The significant positive coefficient for the sine intercept shifter in the MDCE model for both rivers implies that respondents who live north of the river site to be restored, where $\sin(90^\circ) = 1$, are more likely to choose further restoration measures than respondents who live south of the river site, where $\sin(270^\circ) = -1$. A possible explanation for these significant directional effects is that the travel time is likely to be longer if the river sites are approached from the south than from the north due to heavier road traffic congestion, especially around the city of Zurich, which is located south of the two study sites. Moreover, for respondents living north and west of the river Thur the number of substitute sites on which their tax money could be spent is more limited in comparison with other respondents in the sample because of their proximity to the German border.

Finally, the coefficient estimates for the distance to the alternative river site in the two SDCE choice models in Table 3 indicate that there are significant substitution effects between the Thur and Töss river sites. This implies that respondents in the SDCEs consider the same ecosystem services and biodiversity provided by the river Thur as substitutes for those provided by the river Töss, and vice versa. However, in the MDCE choice model, the distance to the alternative river site is significant only in the utility function for the Thur and has a negative sign, suggesting complementarity between the two sites. Hence, there seems to be a shift in the way respondents perceive the alternative river site depending

on whether they are asked to value the restoration of the two river sites independently or simultaneously. This also seems to apply to the river Töss, where the significant substitution effect detected in the SDCE becomes insignificant in the MDCE. The third hypothesis is, therefore, rejected for the SDCEs in favor of the alternative hypothesis that the two rivers provide substitute services. The third hypothesis is also rejected in the utility function for the river Thur in the MDCE, but here respondents perceive the ecosystem services and biodiversity provided by the river Töss as complementary to those provided at the Thur. The third hypothesis cannot be rejected for the river Töss in the MDCE.

Testing Differences in Preference Parameters

In order to test the fourth and final hypothesis that preferences of respondents for river restoration differ between the two sites, Swait and Louviere's (1993) test procedure is applied to the two SDCEs. The test results for both the attribute-only and the extended models show that the null hypothesis of equality of preference parameters between the two SDCE samples is convincingly rejected at the 1% level (the outcome of the likelihood-ratio test statistic is 49.03 with 17 degrees of freedom for the attribute-only models, and 77.72 with 27 degrees of freedom for the extended models). These findings confirm our fourth hypothesis, as they indicate that choice behavior differs significantly between the two SDCE samples, and respondents have significantly different preferences for the restoration of the rivers Thur and Töss. Consequently, the utility functions between the two rivers are not transferable. However, due to the confoundedness of preference and scale parameters, it is impossible to conclude whether these differences are a result of differences in the preference parameters only or in both preference and scale parameters (Louviere, Hensher, and Swait 2000).

The Wald test is used to verify whether the preference parameters in the MDCE model differ between the site-specific utility functions, that is, the first restoration alternative for the river Thur and the second for the river Töss. The test results for both the attri-

bute-only and the extended models show that the two ASCs are significantly different, thus justifying the use of the MDCE. Although respondents in the MDCE sample in general favor restoration of the river Töss, they value the option to walk along the river Thur significantly more than they do for the river Töss. Significant differences are also detected for the price attribute. Respondents who prefer restoration of the river Thur are more sensitive to price increases than respondents who express a preference for restoring the river Töss. The Wald test outcomes for both the attribute-only and the extended MDCE models, therefore, also confirm the fourth hypothesis and suggest nontransferability of the utility functions between the two rivers. Including control for spatial preference heterogeneity for the valued sites, as well as their substitutes, in the extended models diminishes differences between the utility functions compared to the attribute-only models, but the utility function transfer is still deemed inappropriate. Hence, spatial trend variables help to further explain the variation in public preferences for river restoration, but insufficiently to enable transferability of welfare functions between the two river sites.

Table 5 summarizes the various hypotheses tested in the paper and the outcomes of the test procedures across different samples.

6. Discussion and Conclusions

This study developed and applied a novel research design in the discrete choice modeling literature to assess the impact of alternative framing of substitute sites on substitution effects, choice behavior, and welfare estimates. Only a very few studies exist that focus on framing effects when expanding the range of substitute goods in an MDCE. An important distinguishing feature of the study presented here is that SDCEs are implemented as control groups to formally test the effect of framing of substitutes. Spatial preference heterogeneity is, furthermore, accounted for through the inclusion of various spatial trend variables in the choice models. These include distances, directions, and combinations thereof to the

Table 5
Results of Hypothesis Testing

Hypothesis	Thur SDCE	Töss SDCE	Thur MDCE	Töss MDCE	Outcome
$H_0^{1a} : MWTP_{Thur}^{SDCE} = MWTP_{Thur}^{MDCE}$			Rejected for 4 out of 5 nonmonetary choice attributes		
$H_0^{1b} : MWTP_{Töss}^{SDCE} = MWTP_{Töss}^{MDCE}$			Rejected for 3 out of 5 nonmonetary choice attributes		
$H_0^{2a} : \beta_{Distance\ to\ valued\ river\ site} < 0$	Cannot be rejected	Cannot be rejected	Rejected	Cannot be rejected	Cannot be rejected
$H_0^{2b} : \beta_{Longitudinal\ distance\ to\ valued\ river\ site} < 0$	Rejected	Rejected	Cannot be rejected	Cannot be rejected	Cannot be rejected
$H_0^{2c} : \beta_{Latitudinal\ distance\ to\ valued\ river\ site} < 0$	Cannot be rejected	Rejected	Rejected	Cannot be rejected	Cannot be rejected
$H_0^3 : \beta_{Distance\ to\ alternative\ river\ site} = 0$	Rejected in favor of H_{a1}^3 (substitution)	Rejected in favor of H_{a1}^3 (substitution)	Rejected in favor of H_{a2}^3 (complementarity)	Rejected in favor of H_{a2}^3 (complementarity)	Cannot be rejected
$H_0^4 : \beta_{Thur} \neq \beta_{Töss}$	Cannot be rejected for SDCEs	Cannot be rejected for MDCE			Cannot be rejected

Note: MDCE, multiple-site discrete choice experiment; SDCE, single-site discrete choice experiment.

river sites under valuation and alternative river sites.

Framing of a DCE as a valuation task of the same services provided by single or multiple sites significantly affects the level of substitutability between the sites, choice behavior, and WTP values. Local residents and visitors perceive the identical ecosystem services and biodiversity provided by the same site differently, depending on whether the site is valued in isolation or conjointly in the same choice set with other similar sites. Therefore, merely reminding respondents about substitutes, as is typically done in the SP literature, or including them in a choice set leads to significantly different DCE results.

The MDCE generates higher WTP values than the SDCEs, suggesting that complementarity exists between the two river sites as marginal utility increases with additional units of similar services (Hailu, Adamowicz, and Boxall 2000). This outcome could, however, also suggest that some respondents behaved more strategically in the MDCE than in the SDCEs. Our study design could have potentially led to differences in incentive compatibility associated with strategic behavior between SDCE and MDCE versions. Strategic behavior is likely to be more prominent for choices between alternative provision levels of a public good at two different locations than for choices between alternative provision levels of the same public good at a single location. Since respondents in the MDCE were informed that only one of the river sites would be restored, they may have assumed that if more people select the alternative associated with their preferred river site, the more likely it is that this site will be restored. Therefore, a portion of the sample may have had an incentive to systematically choose the alternative associated with their preferred river mainly out of concern that if too many people chose the other river, that other river would be restored and they would not benefit, or would benefit to a lesser extent. Although we do not find an indication for strategic behavior among respondents who received the MDCE version, as only 2.5% of them consistently chose the same river site in all six choice tasks, Carson and Groves (2007) deem that any response format with more than two alternatives is gen-

erally not incentive compatible. They argue that respondents tend to reduce multinomial choices to a binary choice between the two alternatives that they believe will most likely be chosen, independently of their preferred alternative. Nevertheless, DCEs with three alternatives, like those used in this study (i.e., two hypothetical alternatives and one status quo alternative), are common practice in the valuation literature (e.g., Concu 2007; Campbell, Hutchinson, and Scarpa 2009; Jørgensen et al. 2013). Moreover, when put to the test, Collins and Vossler (2009) find no evidence that DCEs with three alternatives are less incentive-compatible than those with two alternatives. The decision on whether to include the relevant range of substitutes in a choice set may ultimately imply a trade-off between possibly introducing a strategic bias or a bias caused by ignoring relevant substitute goods in respondents' choice process. More research into this area seems needed.

Significant differences in welfare estimates are found between the two river sites irrespective of the framing of substitute sites. Local residents and visitors value the same ecosystem services and biodiversity differently across the two rivers, meaning that benefit estimates for the same service provided by one site cannot simply be transferred to another site. Lizin et al. (2016) show that taking substitution effects into account decreases the relative transfer error. Similarly, our study demonstrates that including further control for spatial preference heterogeneity reduces differences between the river-specific utility functions and derived WTP welfare estimates. Although this does not occur to the extent that the utility functions become transferable across sites, possibly due to several differences between the two rivers, it implies that accounting for distance to substitute sites and heterogeneity in distance decay has the potential to considerably reduce transfer errors and enable transferability of site-specific utility functions. This has important implications for the practical use of site-specific SP values in benefits transfer that support decision-making based on a benefit-cost analysis.

Only controlling for unidirectional distances to the site of interest, as is common practice in the SP literature, is insufficient

for properly capturing the full range of spatial variation in preferences and WTP values. Since these can vary across the different directions from the study site, owing to the spatial distribution of substitute sites, one needs to control for such directional effects, as also recommended by Cameron (2006) and Schaafsma et al. (2013). Moreover, the spatial distribution of urban centers is another important factor to consider when analyzing spatial preference heterogeneity. We find evidence that distance decay is not homogeneous across urban and rural areas. Failing to control for any of these spatial trend variables may undermine the validity and reliability of using SP results for environmental improvements in the public policy they aim to inform.

Acknowledgments

We thank Amael Paillex, Peter Reichert, and Andri Bryner for their help during the survey design, Rosi Siber for creating the maps, and the team of interviewers who conducted the survey. Marije Schaafsma provided useful advice concerning the econometric analysis of spatial trends. We are also grateful to Jeroen van den Bergh, Julia Martín-Ortega, Matteo Mattmann, and two anonymous reviewers for their valuable comments. The authors are solely responsible for the content of the paper.

References

- Bateman, Ian J., Brett H. Day, Stavros Georgiou, and Iain Lake. 2006. "The Aggregation of Environmental Benefit Values: Welfare Measures, Distance Decay and Total WTP." *Ecological Economics* 60 (2): 450–60.
- Bernhardt, E. S., M. A. Palmer, J. D. Allan, G. Alexander, K. Barnas, S. Brooks, J. Carr, S. Clayton, C. Dahm, J. Follstad-Shah, et al. 2005. "Synthesizing U.S. River Restoration Efforts." *Science* 308 (5722): 636–37.
- Bhat, Chandra R. 2001. "Quasi-random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model." *Transportation Research B* 35 (7): 677–95.
- Brouwer, Roy, and Julia Martín-Ortega. 2012. "Modeling Self-censoring of Polluter Pays Protest Votes in Stated Preference Research to Support Resource Damage Estimations in Environmental Liability." *Resource and Energy Economics* 34 (1): 151–66.
- Brouwer, Roy, and Louis H. G. Slangen. 1998. "Contingent Valuation of the Public Benefits of Agricultural Wildlife Management: The Case of Dutch Peat Meadow Land." *European Review of Agricultural Economics* 25 (1): 53–72.
- Brown, Thomas C., and John W. Duffield. 1995. "Testing Part-Whole Valuation Effects in Contingent Valuation of Instream Flow Protection." *Water Resources Research* 31 (9): 2341–51.
- Cameron, Trudy Ann. 2006. "Directional Heterogeneity in Distance Profiles in Hedonic Property Value Methods." *Journal of Environmental Economics and Management* 51 (1): 26–45.
- Campbell, Danny, W. George Hutchinson, and Riccardo Scarpa. 2009. "Using Choice Experiments to Explore the Spatial Distribution of Willingness to Pay for Rural Landscape Improvements." *Environment and Planning A* 41 (1): 97–111.
- Carson, Richard T., and Theodore Groves. 2007. "Incentive and Informational Properties of Preference Questions." *Environmental and Resource Economics* 37 (1): 181–210.
- Collins, Jill P., and Christian A. Vossler. 2009. "Incentive Compatibility Tests of Choice Experiment Value Elicitation Questions." *Journal of Environmental Economics and Management* 58 (2): 226–35.
- Concu, Giovanni B. 2007. "Investigating Distance Effects on Environmental Values: A Choice Modelling Approach." *Australian Journal of Agricultural and Resource Economics* 51 (2): 175–94.
- Cummings, Ronald G., Philip T. Ganderton, and Thomas McGuckin. 1994. "Substitution Effects in CVM Values." *American Journal of Agricultural Economics* 76 (2): 205–14.
- FOEN (Swiss Federal Office for the Environment). 2011. *Restoration of Rivers and Lakes in the Water Protection Act* [in German]. Available at www.bafu.admin.ch/dokumentation/medieninformation/00962/index.html?lang=de&msg-id=38964 (accessed October 24, 2015).
- Greene, William H. 2002. *Econometric Analysis*, 5th ed. Upper Saddle River, NJ: Prentice Hall.
- Hailu, Atakelty, Wiktor L. Adamowicz, and Peter C. Boxall. 2000. "Complements, Substitutes, Budget Constraints and Valuation." *Environmental and Resource Economics* 16 (1): 51–68.
- Hanley, Nick, Felix Schläpfer, and James Spurgeon. 2003. "Aggregating the Benefits of Environmental Improvements: Distance-Decay

- Functions for Use and Non-use Values." *Journal of Environmental Management* 68 (3): 297–304.
- Hensher, David A., John M. Rose, and William H. Greene. 2005. *Applied Choice Analysis: A Primer*. Cambridge, UK: Cambridge University Press.
- Hoehn, John P., and John B. Loomis. 1993. "Substitution Effects in the Valuation of Multiple Environmental Programs." *Journal of Environmental Economics and Management* 25 (1): 56–75.
- Hynes, Stephen, Nick Hanley, and Riccardo Scarpa. 2008. "Effects on Welfare Measures of Alternative Means of Accounting for Preference Heterogeneity in Recreational Demand Models." *American Journal of Agricultural Economics* 90 (4): 1011–27.
- Johnston, Robert J., and Mahesh Ramachandran. 2014. "Modeling Spatial Patchiness and Hot Spots in Stated Preference Willingness to Pay." *Environmental and Resource Economics* 59 (3): 363–87.
- Jørgensen, Sisse Liv, Søren Bøye Olsen, Jacob Ladenburg, Louise Martinsen, Stig Roar Svenningsen, and Berit Hasler. 2013. "Spatially Induced Disparities in Users' and Non-users' WTP for Water Quality Improvements: Testing the Effect of Multiple Substitutes and Distance Decay." *Ecological Economics* 92: 58–66.
- Kurth, Anne-Marie, and Mario Schirmer. 2014. "Thirty Years of River Restoration in Switzerland: Implemented Measures and Lessons Learned." *Environmental Earth Sciences* 72 (6): 2065–79.
- Lancaster, Kelvin J. 1966. "A New Approach to Consumer Theory." *Journal of Political Economy* 74 (2): 132–57.
- Lizin, S., R. Brouwer, I. Liekens, and S. Broeckx. 2016. "Accounting for Substitution and Spatial Heterogeneity in a Labelled Choice Experiment." *Journal of Environmental Management* 181: 289–97.
- Loomis, John B. 2000. "Vertically Summing Public Good Demand Curves: An Empirical Comparison of Economic versus Political Jurisdictions." *Land Economics* 76 (2): 312–21.
- Louviere, Jordan J., David A. Hensher, and Joffre D. Swait. 2000. *Stated Choice Methods: Analysis and Application*. Cambridge, UK: Cambridge University Press.
- Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green. 1995. *Microeconomic Theory*. New York: Oxford University Press.
- McFadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, ed. Paul Zarembka, 105–42. New York: Academic Press.
- National Oceanic and Atmospheric Administration (NOAA). 1993. "Report of the NOAA Panel on Contingent Valuation." *Federal Register* 58 (10): 4601–14.
- Neill, Helen R. 1995. "The Context for Substitutes in CVM Studies: Some Empirical Observations." *Journal of Environmental Economics and Management* 29 (3): 393–97.
- Paillex, Amael, Nele Schuwirth, Armin W. Lorenz, Kathrin Januschke, Armin Peter, and Peter Reichert. 2017. "Integrating and Extending Ecological River Assessment: Concept and Test with Two Restoration Projects." *Ecological Indicators* 72: 131–41.
- Palmer, Margaret A., Holly L. Menninger, and Emily Bernhardt. 2010. "River Restoration, Habitat Heterogeneity and Biodiversity: A Failure of Theory or Practice?" *Freshwater Biology* 55 (1): 205–22.
- Pate, Jennifer, and John Loomis. 1997. "The Effect of Distance on Willingness to Pay Values: A Case Study of Wetlands and Salmon in California." *Ecological Economics* 20 (3): 199–207.
- Poe, Gregory L., Kelly L. Giraud, and John B. Loomis. 2005. "Computational Methods for Measuring the Difference of Empirical Distributions." *American Journal of Agricultural Economics* 87 (2): 353–65.
- Revelt, David, and Kenneth Train. 1998. "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level." *Review of Economics and Statistics* 80 (4): 647–57.
- Rolfe, John, Jeff Bennett, and Jordan Louviere. 2002. "Stated Values and Reminders of Substitute Goods: Testing for Framing Effects with Choice Modelling." *Australian Journal of Agricultural and Resource Economics* 46 (1): 1–20.
- Rolfe, John, and Jill Windle. 2012. "Distance Decay Functions for Iconic Assets: Assessing National Values to Protect the Health of the Great Barrier Reef in Australia." *Environmental and Resource Economics* 53 (3): 347–65.
- Rose, John M., Michiel C. J. Bliemer, David A. Hensher, and Andrew T. Collins. 2008. "Designing Efficient Stated Choice Experiments in the Presence of Reference Alternatives." *Transportation Research Part B* 42 (4): 395–406.
- Schaafsma, Marije, and Roy Brouwer. 2013. "Testing Geographical Framing and Substitution Effects in Spatial Choice Experiments." *Journal of Choice Modelling* 8: 32–48.

- Schaafsma, Marije, Roy Brouwer, Alison Gilbert, Jeroen van den Bergh, and Alfred Wagtendonk. 2013. Estimation of Distance-Decay Functions to Account for Substitution and Spatial Heterogeneity in Stated Preference Research." *Land Economics* 89 (3): 514–37.
- Schaafsma, Marije, Roy Brouwer, and John Rose. 2012. "Directional Heterogeneity in WTP Models for Environmental Valuation." *Ecological Economics* 79: 21–31.
- Swait, Joffre, and Jordan Louviere. 1993. "The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models." *Journal of Marketing Research* 30 (3): 305–14.
- Train, Kenneth E. 2003. *Discrete Choice Models with Simulation*. Cambridge, UK: Cambridge University Press.
- Varian, Hal R. 2014. *Intermediate Microeconomics: A Modern Approach*, 9th ed. New York: W. W. Norton and Company.
- Woolsey, Sharon, Florence Capelli, Tom Gonser, Eduard Hoehn, Markus Hostmann, Berit Junker, Achim Paetzold, Christian Roulier, Steffen Schweizer, Scott D. Tiegs, et al. 2007. "A Strategy to Assess River Restoration Success." *Freshwater Biology* 52 (4): 752–69.
- Woolsey, Sharon, Christine Weber, Tom Gonser, Eduard Hoehn, Markus Hostmann, Berit Junker, Christian Roulier, Steffen Schweizer, Scott D. Tiegs, Klement Tockner, and Armin Peter. 2005. *Handbook for Evaluating Rehabilitation Projects in Rivers and Streams*. A publication by the Rhone-Thur Project. Zurich: Swiss Federal Institute for Aquatic Science and Technology (Eawag), Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Laboratory of Hydraulic Constructions (LCH, EPFL Lausanne), and Laboratory of Hydraulics, Hydrology and Glaciology (VAW, ETH Zurich).