



Public preferences for heritage conservation strategies: a choice modelling approach

David Throsby¹ · Anita Zednik²  · Jorge E. Araña³

Received: 14 January 2019 / Accepted: 2 January 2021
© The Author(s) 2021

Abstract

Studies aiming at valuing cultural and natural heritage projects are often focussed on one or only a few sites, whereas planning decisions concerning the allocation of public funds to heritage conservation deal with classes of heritage rather than single sites. In addition, such planning decisions are almost always concerned with non-monetary values that need to be incorporated into assessment procedures if the total value of alternative strategies is to be estimated. In this paper, we put forward and estimate models to address both of these issues within a choice-modelling framework. The method is developed in the context of conservation of a particular class of cultural heritage, namely major historic buildings in a city or country. We report results from a discrete choice experiment to assess public preferences in which the choices are alternative conservation programs and the attributes are dimensions of the programs' cultural and economic value. The model is estimated from survey data using several flexible econometric specifications. We show that the methods developed can be used to obtain robust estimates of the economic value of this category of buildings. We also find a significant contribution of all aspects of cultural value to the formation of conservation preferences and the public's willingness to pay.

This research was funded by a grant from the Commonwealth Environment Research Facilities of the Australian Department of Environment, Water, Heritage and the Arts. The Grant was administered by the Australian National University, Contract No. 17162.

✉ Anita Zednik
anita.zednik@wu.ac.at

David Throsby
david.throsby@mq.edu.au

Jorge E. Araña
jorge.arana@ulpgc.es

¹ Department of Economics, Macquarie University, Sydney, Australia

² Institute for Markets and Strategy, WU Vienna, Vienna, Austria

³ Departamento de Análisis Económico Aplicado, Univesidad de las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain

Keywords Heritage conservation · Public preferences · Choice experiments

1 Introduction

In a paper published in the early 1990s, Alan Peacock argued that public preferences should be taken into account in decisions concerning the conservation of cultural heritage assets, especially when public funding was involved (Peacock 1994). Since then, interest has grown in the assessment of public attitudes to heritage conservation through the application of preference evaluation methodologies such as contingent valuation and discrete choice modelling (Rolfe and Windle 2003; Alberini and Longo 2006; Choi et al. 2010; Tuan and Navrud 2007, 2008; Apostolakis and Jaffry 2005b). In this wide range of studies, two significant issues have emerged which to date have been insufficiently explored. The first relates to the object of assessment. Almost all the studies that have been undertaken relate to a specific building, archaeological site, etc., yet heritage policy decisions are often concerned with the allocation of funds across a range of heritage projects rather than simply to one particular case. Policymakers in heritage administration are routinely faced with decisions requiring an understanding of conservation preferences for a generalised class of buildings or sites rather than for a single specific project (Deodhar 2004; Provins et al. 2008; Lazrak 2009).

The second issue that remains inadequately investigated is the influence of cultural motivations on the formation of people's preferences for different types of heritage, and in particular the relationship between economic and cultural assessments in an individual's personal evaluation. This question can be located in the context of the distinction between economic and cultural value in the assessment of the demand both for cultural goods generally (Throsby 2001; Hutter and Throsby 2008; Throsby and Zednik 2014; Crossick and Kaszynska 2016) and for heritage in particular (Rizzo and Throsby 2006; Throsby 2013).

The objective of this paper is to deal with both of these issues: to provide information on public preferences for a class or category of cultural heritage rather than a single building; and to increase understanding of how cultural values translate into economic preferences. To address the first issue, we construct and implement a choice experiment in which a random sample of respondents drawn from the general public is asked to choose between conservation programs for buildings characterised as being within a particular class of heritage, namely the class of major historic buildings, i.e. those buildings that are included on an official list or register compiled at city, state or national level, and as such are likely to be recognisable to members of the public. In regard to the second issue, the attributes that we include in our investigation are drawn from the presumed dimensions of this category's cultural value. Relevant to both issues, the modelling is framed in terms of the costs of alternative conservation strategies, enabling us to analyse the links between economic and cultural value attributes for the designated heritage class as perceived by respondents.

Researchers interested in exploring preferences for any type of good or service have a choice between revealed preference (RP) and stated preference (SP)

methods. RP studies analyse in retrospect actual behaviour rather than hypothetical choices, the advantage being that there is no need for assumptions about how intentions translate into behaviour. However, there are many limitations affecting the use of RP data (Louviere et al. 2000). For example, the range of variation in explanatory variables in RP data is frequently limited, such that behavioural changes in response to changes in these variables may be difficult to predict. Moreover, explanatory variables tend to be highly collinear in real markets. But most importantly for our study, RP data simply do not exist for goods and services that are not traded on any market, such as public goods, or policy scenarios as yet untried. Thus, in the absence of RP data, the use of SP methods was the only way we could probe the nature of the cultural valuations for heritage under study. Fortunately, the methodologies involved have been widely tested and refined, and we have no reason to doubt their efficacy for our study.

The structure of the paper is as follows. In the next section, we consider the theoretical background to the application of discrete choice methods in the cultural heritage field. The following section includes a review of econometric specifications that have been proposed to analyse choice data sets. We then outline the choice experiment, beginning with a description of cultural heritage categories, the attributes that characterise the heritage class under study, the hypotheses relating to their relative influence, and details of the experimental design. The next section discusses the econometric approach, with an outline of various possible models. We then present results for the preferred model, a mixture-of-normals multinomial logit (MM-MNL) model in preference space, and willingness-to-pay (WTP) estimates for all attributes. After a discussion of the results, the final section puts forward some conclusions, caveats and proposals for future research. An [Appendix](#) to the paper sets out econometric specifications of all the models used.

2 Theoretical background

The theoretical foundations for this paper are derived from the theory of preference formation in economics. Among approaches to the empirical investigation of consumer preferences, discrete choice experiments have found an important place. Choice experiments originally arose from conjoint analysis and have been used in transportation, marketing and psychology (see for example Louviere and Woodworth 1983; Hensher 1994). Conjoint analysis typically asks respondents to rank or rate goods or attributes, whereas in choice experiments, respondents choose from alternative bundles of attributes, thus making their choices consistent with random utility theory (Thurstone 1927; McFadden 1974; Ben-Akiva et al. 1985). Choice experiments are also an application of the characteristics theory of value (Lancaster 1966), since in choice experiments goods are broken into attributes, where one of these attributes is usually price. Choice experiments have become an important and established methodology in environmental economics (Hanley et al. 1998; Adamowicz et al. 2014), since passive-use (non-use) value is of such importance for environmental goods. Adamowicz et al. (1998) were among the first to estimate non-use value for environmental resources; they undertook choice experiments for

a particular site of old-growth forests in Alberta, Canada, where caribou, an endangered species, live. Nowadays, choice experiments are employed in a wide range of fields when revealed preference data are not available or as a supplement for revealed preference data (Rolfe et al. 2000; Campbell et al. 2009; Walker and Marquis-Kyle 2004; Strazzera et al. 2010; Mas and Pallais 2017).

In cultural economics, discrete choice experiments have been applied to several areas, for example to forecast demand for a cultural event (Louviere and Hensher 1983), to investigate theatre demand (Willis and Snowball 2009; Grisolia and Willis 2011, 2012), or to study museum attendance (Maddison and Foster 2003; Jaffry and Apostolakis 2011). In the particular field of cultural heritage research, researchers (Morey and Rossmann 2003; Morey et al. 2002) conducted a discrete choice experiment to estimate the benefits of reducing acid deposition injuries to 100 marble monuments in Washington, DC. Choi et al. (2010) conducted a choice experiment on the value of Old Parliament House, a national heritage site in Canberra, Australia, as well as on the various services it provides to the public. Rolfe and Windle (2003) employed a choice experiment to assess the trade-off between the protection of Aboriginal heritage sites and waterways development in the face of increased irrigation demands in the Fitzroy Basin in Australia. Other examples of choice experiments for cultural heritage come from Taiwan (Chen and Chen 2012), New Zealand (Miller et al. 2015), Greece (Apostolakis and Jaffry 2005b, a), Portugal (Lourenço-Gomes et al. 2014) and Italy (Mazzanti 2003). In Ireland, the hedonic pricing valuation method was applied to estimate the value of cultural heritage in the housing market (Moro et al. 2013).

A particular focus of the present study is on the notion of cultural value, a concept that has crystallised in recent years as a form of value distinct from conventional interpretations of economic value (Angelini and Castellani 2019). In the context of cultural heritage, this duality of value—cultural and economic—derives from the interpretation of heritage items as cultural capital assets (Throsby 1999; Rizzo and Throsby 2006; Apostolakis and Jaffry 2007), defined as capital goods that embody or yield cultural value in addition to whatever economic value they possess. It is understood that economic value, whether measured as direct use value or willingness to pay for non-use demand, is expressible in monetary terms, whereas cultural value is characterised by multidimensionality and has no single unit of account. The latter characteristic places cultural value outside the framework of pecuniary value inherent in neoclassical economics. To operationalise cultural value for any cultural good or service, it can be deconstructed into its constituent elements, identified in general terms as relating to the aesthetic and symbolic properties of the good or service in question.

Cultural value is of particular relevance to heritage, where criteria to determine the cultural significance of heritage buildings and sites are used by heritage professionals in evaluating the relative importance of historic buildings, archaeological sites, landscapes and so on (Avrami et al. 2000). For example, a standard procedure for evaluating and managing these sorts of heritage items is put forward in the so-called Burra Charter (Walker and Marquis-Kyle 2004), and UNESCO specifies similar criteria for determination of the significance of items of heritage to be included in the World Heritage List of heritage of “universal human value”. A variety of

possible dimensions or characteristics can be proposed to define cultural value in particular circumstances; in the present study, we chose aesthetic, historical, social and architectural values as the most appropriate for our purposes, as explained further in the following Sect. 3.

In economic analyses of heritage, this multidimensional nature of value is likely to be recognised by incorporation of some assessment of cultural significance alongside the standard economic measurements of use and non-use value (Mazzanti 2002; Mason 2005; Clark and Maer 2008; Mason 2008). In these studies, assessment of significance is a separate issue; the approach we develop in this paper incorporates assessment of economic and cultural value into a single integrated methodology.

3 The choice experiment

3.1 Pre-testing of concepts

As noted above, the choice experiment was set up to study public preferences for alternative conservation programs applied to different categories of built heritage. The first stage of the study involved two focus groups to test the extent to which participants recognised different categories of heritage, and to explore their perceptions of the cultural significance of heritage buildings. Five categories of heritage were put to participants:

- Major historic buildings or groups of buildings of national or state significance
- Local historic buildings or groups of buildings, of significance in local communities
- Residential houses or groups of houses from earlier times
- Rural landscapes or townscapes with historic structures
- Sites or landscapes of indigenous significance

The categories were described in detail and photographs of typical buildings in each category were distributed for discussion. The focus group work confirmed that all five categories could be readily distinguished by those participating using the wording and picture aids that we went on to employ in the final survey.

To assess perceptions of cultural value, we specified a number of criteria related to cultural significance and invited participants in the focus groups to consider them and to nominate any further characteristics of heritage buildings that they regarded as important. Analysis of the results of these procedures enabled us to identify the criteria that were most significant in respondents' minds. We used these as the basis for specifying the cultural value attributes included in the choice experiment, as described further as follows.

3.2 The survey questionnaire

The questionnaire began by introducing respondents to the concept of cultural heritage; photographs of heritage buildings in the above categories drawn from the Australian Heritage Database were presented and respondents were asked a series of questions to gauge their views on the importance of heritage for the Australian community and their attitudes to public funding to support heritage conservation. They were then told that the rest of the survey would focus on one specific type of heritage, namely major “iconic” cultural heritage buildings, the category most likely to be recognisable to survey respondents. Some such buildings may be of international significance—unique Australian heritage icons such as the Sydney Opera House that are known to many people around the world. At a more local level, citizens may be aware of major historic buildings located in their own city or town, or in other national or regional cities or towns. As a reminder, we showed them again photographs of the major cultural heritage category. Participants were then shown a screen that stated different points expressing why some people say that major historical buildings are valuable to the Australian community, and another series of points expressing the opposite view. The purpose of these procedures was to give respondents a sense of what the category “major historic buildings” comprises and to expose them to both positive and negative opinions about heritage so that they might be able to crystallise their own views.

The next section of the questionnaire described the nature of conservation as it applies to heritage buildings. We outlined the purposes that a one-off levy for heritage preservation could be used for, such as: protecting buildings from demolition, major alterations, re-development or neglect; providing more money for restoration and maintenance; and opening up government-owned sites to the public. By doing so, we implicitly defined the base “no conservation” condition, since these efforts would not take place without such funding.

As noted above, we chose four dimensions of cultural value to include as attributes of the major historical buildings category in the survey. The rationale for presenting these attributes to respondents was as follows:

- *Aesthetic value* In specifying beauty as an attribute of heritage buildings in this study, we did not refer to a respondent’s individual and possibly idiosyncratic personal evaluation, but rather we framed this attribute in terms of a general or average judgement by using the phrase “noted for their beautiful appearance”. In this way, we were able to assess the extent to which aesthetic quality of heritage figures as important in the respondent’s opinion.
- *Historical value* The age of a building is an objective characteristic that contributes importantly to its cultural value. It can be argued that the historical value of heritage buildings is significant because their history provides a narrative context for interpretation of their importance, and it allows observers to think about connections between present and the past.
- *Social value* Heritage sites can play an important role in community life, for example as a location for community engagement and social interaction. They may also have significance as symbols of national or local identity, contributing

Table 1 Attributes and attribute levels of major cultural heritage buildings used in the choice experiment

Attributes	Attribute levels
Aesthetic value	Buildings that are not particularly noted for their beautiful appearance Buildings that are especially noted for their beautiful appearance
Social value	Buildings that do not play any particular role in community life Buildings that play an extremely important role in community life
Architectural value	Buildings that are of no particular architectural importance Buildings that are important examples of a particular architectural style
Historic value	Buildings that date from various historic times Buildings that date from Australian colonial times, i.e. 1788–1899 Buildings that date from the early 20th century, i.e. 1900–1940 Buildings that date from the post-war period, i.e. 1940s and later
Levy	\$25 \$50 \$75 \$100

to social cohesion and unity. We gather these various dimensions of social value into an attribute concerning the role that the heritage plays (or does not play) in community life.

- *Architectural value* Heritage buildings may have significance from an architectural viewpoint on account of their originality, their influence on architectural trends, or their typicality of a particular period. An objective assessment of architectural significance is a matter usually left to experts rather than the public. Hence, in specifying this attribute, we referred to what a respondent might see as a collective judgement about architectural importance, and thus we can assess the extent to which such a characteristic of historic heritage features as an item in the formation of the individual's preferences.

Table 1 shows the levels of the attributes included in the choice experiment. Three of the cultural value attributes—*aesthetic*, *social* and *architectural*—were specified as binary variables. The fourth—*historical value*—was divided into four attribute levels corresponding to four different eras in history, as can be seen in the table.¹ The setting for the experiment was presented to participants as one in which the government was considering initiating a program of conservation for a number of major historic buildings, to be financed via a one-off levy that people would pay via their

¹ For the estimation, we effects-coded the historic periods (four levels) and used “date from various historic times” as base. The three attributes beautiful appearance, social role and architectural significance were dummy coded, where 1 indicates that this aspect of cultural value is particularly strong and 0 indicates that the cultural value aspect is not particularly strong in a particular heritage conservation program.

Table 2 Example of a choice task

	Program 1 focuses on major historic buildings that	Program 2 focuses on major historic buildings that	Program 3 focuses on major historic buildings that	Program 4 focuses on major historic buildings that
Historic importance	... date from colonial times, i.e. 1788 to 1899	... date from various historic times	... date from various historic times	date from the early 20th century, i.e. 1900 to 1940
Beauty	... are not particularly noted for their beautiful appearance	... are not particularly noted for their beautiful appearance	... are not particularly noted for their beautiful appearance	... are especially noted for their beautiful appearance
Social role	... play an extremely important role in community life	... play an extremely important role in community life	... play an extremely important role in community life	... do not play any particular role in community life
Architectural significance	... are not of particular architectural importance	... are important examples of a particular architectural style	... are not of particular architectural importance	... are important examples of a particular architectural style
Price	\$25	\$75	\$50	\$25

Program 1	Program 2	Program 3	Program 4	None
-----------	-----------	-----------	-----------	------

taxes. Thus, a fifth attribute was specified to represent the price that would be levied on an individual's tax as a contribution towards the cost of the program. During the focus groups, open-ended questions concerning willingness to pay for different combinations of attributes were presented to participants to obtain information useful for designing the cost attribute levels.² Values mentioned lay in the range between \$0 and \$100. It was indicated that different possible programs were being considered, of which only one could be implemented. Each of the programs would be focussed on heritage buildings with a different combination of qualities such as their beauty, their age, their architectural significance, their social significance and their tax price. Respondents were asked which one they would choose. The experiment was then repeated with the different programs specified with different sets of attributes.³

Conservation efforts necessarily represent a trade-off between many different subgroups of buildings within a heritage category, each of which might be more attractive in some cultural value aspects whilst being less attractive in others. The trade-off between many alternatives for conservations is part of the problem of cultural heritage protection. In addition, we take advantage of the statistical gains of larger choice set sizes. Table 2 shows an example choice set that was shown to participants before the actual choice experiment.

² An open-ended question in this context might read: 'How much would you be willing to pay for this specific combination of programs to be implemented?'

³ The discrete choice task was introduced with the following instructions: 'On the next screens you will see four possible conservation programs that focus on major historical buildings with different characteristics. The programs also cost different amounts, measured as a one-off levy every Australian would pay via the tax system to provide the funds. The proceeds of this levy would be managed through a heritage trust fund. The levy would be set up in a way that if the majority of people agreed to pay a particular levy amount, everyone would have to pay the levy'. The instructions above each choice set were: 'Please tick which of the four cultural heritage conservation programs that focus on major historical buildings you would prefer most and actually support, if the four programs where the only ones proposed. If you believe that none of the programs should be supported, please click on "none".'

3.3 Experimental design

A full factorial for the experimental design for the discrete choice task would have led to 128 profiles. We used a 2^7 orthogonal design that gave us eight profiles. We then created the fold-over of that design to generate eight more profiles, a total of 16 profiles. That ensured that all estimated effects could be independently estimated from each other, as well as independently from first-order interactions. We used the 16 profiles in a balanced incomplete block design that led to 20 choice sets, each containing four of the profiles. Balanced incomplete block designs have the advantage that each profile is presented an equal number of times (in our case five times) and presented equally often with every other profile (in our case once). A “none” option was added to each choice set so that participants could pick this option if they did not want to support any of the conservation programs in that choice set. Each respondent was asked to complete all 20 choice sets.⁴

In reaching a decision on choice set size and number of choice sets, researchers typically face a trade-off between learning and fatigue effects (Czajkowski et al. 2014). On one hand, a larger number of scenarios gives gains in statistical efficiency (more information per respondent) and learning (subjects’ understanding of the choices involved). On the other hand, there might be a loss in the quality of responses when more scenarios are presented, since subjects may feel bored or tired (fatigue effects). Following Johnston et al. (2017), and after testing several choice set sizes and number of choices per respondent in focus groups, we observed that a choice set size of four, and 20 choices per respondent, provided high levels of realism and were a good combination of statistical and respondent efficiency.⁵ The sensitivity of results to these design aspects was also tested by implementing a post-survey nonparametric test analysing responses to the first (and last) five tasks, finding that parameters were quite stable.

In order to avoid any potential bias in presenting options in a certain order (for example a left-to-right survey response bias), we presented the four options within each choice set in a random order. We also randomised the order in which the choice sets were shown to each respondent. In addition, immediately before respondents went through the choice sets, they saw the same five images of examples of major heritage buildings again that they had seen earlier, to refresh their minds about the cultural heritage category concerned. A pilot study was undertaken online one month before the main study to verify that respondents understood the tasks and answered the questions as expected, and to determine how much time respondents took.

⁴ Note that the experiment was set up as an unlabelled experiment, meaning that the labels “Program 1”, “Program 2”, etc., do not have any meaning apart from the order in which they appear in the experiment.

⁵ See Scarpa et al. (2007) and Yao et al. (2015) for an example of the importance of combining both dimensions of efficiency in choice experiments.

3.4 Hypotheses

The choice experiment was set to test a series of a priori hypotheses concerning the importance of cultural value elements in determining respondents' preferences, and their willingness to pay for alternative combinations of these attributes. In broad terms, these hypotheses propose that people will prefer conservation programs that focus on: buildings particularly noted for their beauty; older buildings rather than those from more recent times; buildings that play an important role in social life; and buildings of particular architectural significance. We expect that these preferences will be reflected in respondents' willingness to pay as indicated by their responses on the cost of alternative conservation programs. These hypotheses relate to preferences within attributes. In regard to the relative strengths of preferences between the cultural value attributes, there is no particular basis for formulating a testable hypothesis beyond the proposition that beauty, however interpreted, appears to be a consistently positive influence on decision-making in artistic and cultural contexts,⁶ suggesting that aesthetic value in our choice experiment might be expected to have a stronger effect on choice of conservation programs than the other cultural value attributes.

3.5 Survey implementation

The main experiment was conducted online in August 2012, with a sample of 282 respondents drawn from adult residents in the state of New South Wales, Australia. Respondents were sampled from an online panel hosted by a market research company. Descriptive statistics showing the socio-demographic composition of the sample are shown in Table 3. We used stratified sampling to ensure that the sample would reflect the age and gender distribution of the adult (18+) population of the state. Although some time has elapsed between the date of the survey and the present, we note that the setting of the study and the nature of the information sought and hypotheses tested are not time-dependent but relate to more basic attitudes and preferences in the population. Moreover, it is not the intention of the research to provide direct policy advice but rather to throw light on the nature of public preferences for an important and enduring cultural phenomenon.⁷

3.6 The econometric approach

Consider a range of heritage conservation programs $j = 1, 2, \dots, K$. As in the traditional random utility model (McFadden 1974), let us assume that citizen i chooses a single conservation program j from among k mutually exclusive program

⁶ See, for example, the results of a survey of the cultural value of paintings in Throsby and Zednik (2014).

⁷ We discuss the issue of temporal stability of DCE results further in the concluding section of this paper.

Table 3 Demographic characteristics of the sample

Variable	<i>N</i>	%
Gender		
Male	141	50.00%
Female	141	50.00%
Education		
No senior secondary qualification	31	11.00%
Senior secondary qualification or equivalent	68	24.10%
Post-secondary diploma	79	28.00%
Bachelors or postgraduate degree	104	36.90%
Income		
AUD 0–599/week	58	20.60%
AUD 600–999/week	66	23.40%
AUD 1000–1599/week	57	20.20%
More than AUD 1600/week	57	20.20%
Prefer not to say	44	15.60%
Occupation		
Professional	59	20.90%
Managerial, executive	35	12.40%
Clerical	38	13.50%
Skilled/semi-skilled	35	12.40%
Unskilled	14	5.00%
Serviceman/servicewoman	10	3.60%
Student	27	9.60%
Home duties	33	11.70%
Unemployed	11	3.90%
Other	20	7.10%
Age		
Mean	44.3	
Median	43.5	
SD	15.57	
Min	18	
Max	81	
<i>n</i>	282	

alternatives in a specific choice situation t . For a well-behaved preference map,⁸ upon selecting program j , a general indirect conditional utility U_{ijt} takes the following form:

$$U_{ijt} = V_{ijt} + \epsilon_{ijt} = \beta_i' x_{ijt} + \epsilon_{ijt} \quad i = 1, \dots, N; j = 1, 2, \dots, K; \text{ and } t = 1, 2, \dots, T \quad (1)$$

⁸ Represented by a strictly increasing, continuous and strictly quasi-concave utility function.

where V_{ijt} represents the observed systematic portion of utility and is assumed to be a linear, additive function of the levels of the program attributes x_{ijt} , and β_i is a vector of random variables that allow researchers to account for preference heterogeneity in the population. The random error term, ϵ_{ijt} , represents the unobserved portion of the utility function, which is assumed to be independently and identically distributed (IID) over citizens, conservation programs and choice situations.⁹

Following the behavioural decision process proposed by McFadden (1974), all models considered here depart from the straightforward structure that assumes that, if a citizen faces a multi-attribute discrete choice problem, the researcher will observe that citizen i chooses program j^* if, and only if:

$$U_{ij^*t} = [\beta_i' x_{ij^*t} + \epsilon_{ij^*t}] > U_{ijt} = [\beta_i' x_{ijt} + \epsilon_{ijt}] \quad \forall j \neq j^* \quad (2)$$

Under this general specification of the utility function, the probability that a program j^* is chosen by citizen i , in a choice situation t , is specified as follows;

$$P_{ij^*t} = \text{Prob}(\beta_i' x_{ij^*t} + \epsilon_{ij^*t} > \beta_i' x_{ijt} + \epsilon_{ijt}) \quad \forall j \neq j^* \quad (3)$$

The standard specification presented in Eqs. 1–3 is often called a choice model in preference space. However, as has been noted by Johnston et al. (2017) among others, when the goal of the study is to elicit social preferences—as in the application proposed here—estimating the model directly in what has been called willingness-to-pay (WTP) space, rather than estimating a model in preference space and then calculating parameters of interest by reparametrization, has several advantages especially when the interest lies in the full distribution of the WTP and not just in some moments of the distribution (Train and Weeks 2005; Balcombe et al. 2008, 2009; Scarpa et al. 2008; Thiene and Scarpa 2009; Scarpa et al. 2009; Greene and Hensher 2010).

Our model can be defined in the WTP space by taking into account that the WTP for any attribute i of the heritage policy comes from dividing the marginal utility of the attribute evaluated at the inverse utility function represented in Eq. 1 (i.e. β_i) and the marginal utility of money in the same utility function β_M . Therefore, following Train and Weeks (2005), defining $C_i = \frac{\beta_i}{\beta_M}$ allows us to rewrite Eq. 3 as follows

$$P_{ij^*t} = \text{Prob}(C_i x_{ij^*t} + \epsilon_{ij^*t} > C_i x_{ijt} + \epsilon_{ijt}) \quad \forall j \neq j^* \quad (4)$$

Different specifications and assumptions concerning the deterministic (V_{ijt}) and stochastic (ϵ_{ijt}) portions of the utility function lead to alternative econometric models that have been employed to analyse choice data. In order to explore the sensitivity of our results to the choice of model specification (model uncertainty) in this study, we estimated seven of the most popular econometric models employed in analysing choice experiments. They are the multinomial logit model (MNL), the normal mixed

⁹ Note that since U_{ijt} is an indirect conditional utility function, all the determinants of individuals choices that do not differ among programs at each choice occasion (i.e. individuals' socio-demographic, psychological and cultural variables; choice framing effects like complexity, information level, emotional load,...) cannot be included in (1) unless they are interacted with a constant term.

Table 4 Results of the model selection criteria for alternative econometric approaches

	Marginal likelihood	AIC	BIC	CAIC
MM-MNL	- 468.237	513,08	537,06	541,48
G-MNL	- 471.506	520,29	544,67	549,31
T-MIXL	- 490.75	541,48	576,94	570,09
LC	- 546.007	597,61	622,18	624,75
N-MIXL	- 611.158	672,23	694,49	697,59
S-MNL	- 738.717	812,59	842,14	849,52
MNL	- 800.17	888,19	912,19	912,19

logit model (N-MIXL), the theoretically restricted mixed logit model (T-MIXL), the scaled multinomial logit model (S-MNL), the generalized mixed logit model (G-MNL), the latent class model (LC), and the mixture-of-normals mixed multinomial logit model (MM-MNL). A detailed specification of all models is included in an [Appendix](#) to this paper.

4 Results

As mentioned above, we account for potential model uncertainty by estimating all seven of the above-listed models. All models can be compared in terms of statistical goodness-of-fit and elicited WTP distributions for the proposed attributes in the experiment. Turning first to goodness-of-fit, we show in Table 4 the relevant statistics for the different model specifications. Four criteria are used for comparing the models' goodness-of-fit: marginal likelihood (ML); Akaike information criterion (AIC) (Akaike 1974, 1987); Bayesian information criterion (BIC) (Schwarz 1978); and conditional Akaike information criterion (CAIC) (Bozdogan 1987).

Since we observe the MM-MNL model to be the best fit, we report the results for this model in Table 5. We can see that the buildings most favoured for conservation by respondents are those from colonial times, reflecting a view that it is historic rather than more recent heritage that should be the primary focus of public policy. This result is also consistent with a sense that older buildings are those most likely to be in need of conservation. The results also indicate that respondents prefer conservation efforts that target buildings from pre-20th-century times but their enthusiasm is less pronounced for buildings from the post-war period or the late 20th century. Beautiful appearance, social role and architectural significance are all positive and significant. Not surprisingly, the coefficient for levy is negative, i.e. the higher the levy, the less likely respondents are in general to support the conservation program, other things equal.

Preferences are heterogeneous in the population, but since we do not have any particular hypotheses with respect to different preferences for cultural value aspects of major historic buildings, we simply examine the results to identify three classes of respondents, as shown in Table 5. Class 1 might be considered the greatest fans of colonial times, whilst being less influenced by other criteria. Those in class 2 have

Table 5 Results from the MM-MNL model

	MM-MNL		MM-MNL by class		
	est	SE	Class 1	Class 2	Class 3
Date from late 20th century					
<i>Mean</i>	-0.152	0.008	-0.154	-0.141	-0.168
<i>Std dev.</i>	0.076	0.008			
Date from post-war period					
<i>Mean</i>	-0.647	0.007	-0.497	-0.718	-0.669
<i>Std dev.</i>	0.342	0.005			
Date from colonial times					
<i>Mean</i>	0.291	0.006	0.384	0.288	0.237
<i>Std dev.</i>	0.095	0.006			
Beautiful appearance					
<i>Mean</i>	0.634	0.009	0.497	0.691	0.816
<i>Std dev.</i>	0.491	0.007			
Important social role					
<i>Mean</i>	0.896	0.008	0.853	0.91	0.952
<i>Std dev.</i>	0.317	0.006			
Architecturally significant					
<i>Mean</i>	0.873	0.007	0.763	0.927	0.898
<i>Std dev.</i>	0.569	0.007			
Levy					
<i>Mean</i>	-0.009	0.005	-0.02	-0.01	0
<i>Std dev.</i>	0.006	0.004			
Class probability			0.51	0.22	0.27

Estimated average and standard deviations for individual-specific changes in probabilities

a strong dislike for post-war period buildings—for them the social role and architectural value of cultural heritage buildings is particularly important. For class 3, beautiful appearance, social importance and architectural significance are significant criteria and this class is also least influenced by the amount of the levy.

We now turn to the WTP estimates for different cultural value attributes. Table 6 shows the WTP estimates for each cultural value component for the seven estimated models. These WTP estimates can be interpreted as follows: holding everything else constant, participants would on average be willing to pay the stated amount of money if a conservation program of cultural heritage was focussed on buildings with this particular attribute. As can be seen, the results are fairly robust against different specifications.

In light of the above results, we can draw the following conclusions concerning the hypotheses we proposed about the magnitudes of the cultural value effects. Our hypotheses concerning preferences within attributes are confirmed. However, the data do not support our proposition that the aesthetic qualities of buildings would be a stronger

Table 6 Willingness-to-pay estimates of cultural value attributes for major cultural heritage. Results from models in WTP Space (Australian dollars, AUD)

Attributes	MNL	S-MNL	N-MIXL	T-MIXL	G-MNL	LC	MM-MNL
Date from late 20th century	-22.76 [-22.51, -23.24]	-23.17 [-22.65, -23.69]	-25.06 [-24.79, -25.33]	-24.82 [-24.58, -25.06]	-21.69 [-21.46, -21.92]	-25.37 [-25.21, -25.53]	-26.31 [-26.22, -26.39]
Date from post-war	-68.82 [-68.46, -68.18]	-66.71 [-66.19, -66.23]	-65.42 [-65.15, -65.69]	-64.22 [-63.98, -64.46]	-63.55 [-63.35, -63.75]	-68.27 [-65.64, -65.99]	-69.72 [-69.63, -69.89]
Date from colonial times	31.39 [31.05, 31.73]	31.56 [31.05, 32.07]	32.33 [32.15, 32.71]	31.88 [31.62, 32.14]	30.71 30.52, 30.94]	32.13 [31.95, 32.30]	30.58 30.49, 30.66]
Beautiful appearance	71.39 [71.05, 71.73]	71.56 [69.19, 72.41]	72.73 [69.57, 76.01]	71.88 [70.62, 72.17]	70.71 67.52, 73.97]	72.13 [65.99, 79.30]	70.58 70.92, 70.68]
Important social role	104.4 [101.1, 107.7]	100.9 [93.6, 106.2]	108.7 [105.2, 111.44]	101.7 [98.93, 104.41]	103.5 [101.62, 106.03]	97.3 [96.57, 99.03]	103.91 102.38, 104.57]
Architecturally significant	87.39 [84.01, 90.86]	84.09 [81.86, 88.62]	87.34 [85.62, 89.43]	84.17 [83.93, 84.41]	90.84 [87.62, 93.02]	86.73 [80.88, 92.03]	85.96 83.87, 87.47]

Numbers in brackets are the nonparametric approximations of the lower and upper boundary of the 95% confidence intervals

criterion for allocating conservation funds than the other attributes. In fact, we can see that beautiful appearance is ranked only third in all our models estimated, with social role and architectural significance being judged to be more important.

5 Conclusions

The study reported in this paper leads to several important conclusions. Firstly, we have shown that it is possible to investigate preferences for a particular identifiable class of heritage buildings rather than simply for a single case. This result is significant for the formulation of heritage policy in the public sector at local or national levels, since policymakers are often faced with budgetary decisions framed in these terms. It should now be possible to extend this analysis to examine preferences for other categories of built heritage, such as groups of domestic houses in a heritage district, or local community-based heritage assets.

Secondly, from a methodological/analytical perspective, our results demonstrate that both economic value and cultural value can be incorporated into a single integrated procedure in an evaluation of conservation decisions. In this regard, our methodological approach moves beyond choice modelling applications in the heritage field that focus only on the physical characteristics of publicly accessible heritage sites as they affect visitors (facilities, opening hours, etc.). The cultural value attributes in our study address more fundamental intrinsic qualities of heritage properties as drivers of people's preferences, and the results show the relationship between these characteristics and the economic parameters that influence the formulation of heritage policy decisions. Further research is needed to elaborate the components of cultural value in finer detail than we have been able to in this study.

As is always the case in stated-preference studies, our willingness-to-pay estimates cannot be interpreted as precise dollar amounts, but as indicators providing important input into understanding the economic dimensions of the issue under study. There is a long debate as to whether discrete choice experiments or other hypothetical stated-preference studies systematically overestimate WTP and whether they are susceptible to slight changes in experimental design. Whatever stance one takes in this debate, there is no reason to assume that some cultural value attributes are overestimated more than others, so that our study provides valuable information on the relative magnitude of the individual cultural value components as they relate to economic value of major cultural heritage buildings.

Thirdly, our study has some interesting implications in regard to particular aspects of heritage policy. One important and surprising result is that beautiful appearance does not play such an important role in influencing preferences as might be assumed. In this study, we find that architectural and social significance are the most important drivers of WTP. Respondents have a strong preference for some periods (namely colonial buildings) and show less enthusiasm for heritage buildings from more recent times. Another surprising result is that income does not seem to play a significant role in the likelihood of support for heritage conservation efforts. We also could not find any support for the hypothesis that high-income earners are likely to allocate a larger proportion of their income towards heritage conservation

effort than those on lower incomes. One plausible explanation is that it is wealth rather than income that matters, and we only measured income.

In translating these sorts of procedures into practical application in specific cases, there is a challenge in determining how particular cultural value criteria can be represented in terms of rankings or ratings for funding allocation purposes. Some criteria can be more objectively represented than others. So, for example, the assessment of historic and architectural value might be relatively straightforward, but there may be more controversy in regard to social importance, and even more so in judging beautiful appearance.

To conclude, we return to the issue of whose preferences matter in the disbursement of public funds for heritage conservation that was originally raised by Alan Peacock and that we discussed in the introduction to this paper. Ideally, any decision relating to the allocation of funds for heritage conservation must be based on the best and fullest information available. Of course, expert opinion is indispensable in such a context—these are the heritage professionals with the detailed knowledge and expertise to be able to form sound judgements on cultural and architectural significance of buildings or classes of buildings under consideration. But Peacock's argument is also relevant—the public have a right to make their preferences known as an input into such decision-making. Our study provides an illustration as to how these preferences can be objectively assessed and potentially taken into account.

There are several implications for future research arising from this study. Firstly, the amount of visual or textual material provided to respondents is a critical issue in any choice experiment and is likely to be especially difficult in a DCE concerned with a class of buildings rather than a single case. In our study, we assumed that respondents were able to assimilate all the information provided, such that the results we obtained were based on a true understanding of the different classes of heritage under consideration. But this study is the first of its kind, and further research is needed to determine whether alternative ways of conveying information in these situations may be more efficient, especially with regard to the “no conservation” scenario. Likewise, it is important to consider whether results in a study like this one are robust to changes in experimental design, for example in the number of alternatives or choice sets shown to respondents.

A relevant issue in choice experiments aimed at guiding public policy is the potential existence of attribute non-attendance (ANA), i.e. respondents considering just a subset of attributes for some or all of the choice situations. It has been shown that ANA can seriously affect welfare estimations (Scarpa et al. 2009). There are at least three ways of accounting for ANA in choice experiments. Firstly, respondents can be asked briefing or debriefing questions (before or after the choice task). This is often called a stated ANA (SANA) (Balcombe et al. 2016; Caputo et al. 2017). A second possibility consists in indirectly assessing the ANA by inferring it econometrically (Kragt 2013; Scarpa et al. 2009; Caputo et al. 2017). Finally, some authors have proposed the use of physical tests like eye-tracking data (Balcombe et al. 2015). In our study, a formal test for ANA was not implemented, although we did not observe evidence of ANA during the pre-survey design stages or in verbal protocol techniques employed during focus groups. Nevertheless, future research testing for ANA in our context could be very informative.

Finally, our data are derived from a survey conducted some time ago (2012), raising the question of the temporal stability of elicited preferences using choice experiments. Previous evidence shows that whilst there may be some temporal variations in public preferences elicited in such experiments, differences lie mostly in the marginal utility of income (affecting unconditional WTP), whilst the relative importance of attributes and levels remains stable (Liebe et al. 2012; Schaafsma et al. 2014; Rigby et al. 2016). For our study, there is no obvious reason why the preferences under consideration would have changed in any particular direction in the period since the fieldwork was carried out, and the results do not suggest any such problems. However, the stability of results in future studies of conservation policy using data collected in some previous period remains a matter worthy of further investigation.

Appendix: Alternative specifications of the econometric model

M1. Multinomial logit model (MNL)

The simplest specification of the random utility model is the multinomial logit model (MNL) (McFadden 1974). The indirect conditional utility represented in (equation 1) presents two main assumptions in MNL. First, it assumes homogeneity in the observed part of the utility functions, that is, $\beta_i = \beta; \forall i = 1, \dots, N$. And second, it assumes that the error term, ε_{ijt} is IID distributed with a Type 1 extreme value distribution. Under this specification, the probability of observing a citizen choosing a specific program in any choice situation (equation 3) comes from the following expression:

$$P_{ij^*t} = \exp(\beta' X_{ij^*t}) \sum_{j=1}^k \exp(\beta' X_{ijt}); \forall j \neq j^* \quad (5)$$

In the MNL, the variance of the error term is $\sigma^2\pi^2/6$, where σ is the scale parameter. In order to ensure parameter identification, we follow the standard choice of normalising σ to 1, which results in a variance of the error term $\pi^2/6$.

M2. Normal mixed logit models (N-MIXL)

Several flexible econometric specifications have been proposed aimed at relaxing the restricted assumptions of conventional MNL. The most extensively used option for analysing choice data comes from the use of what can be termed uncorrelated normal mixed logit model (N-MIXL). As in any mixed logit model (MIXL), this specification implicitly accounts for unobserved citizen preference heterogeneity in the sampled population by assuming that β_i is a collection of variables that are independent and drawn from a specific statistical distribution. Thus, the indirect conditional utility in this model is

$$U_{ijt} = (\beta + \eta_i)' X_{ijt} + \varepsilon_{ijt} \quad (6)$$

where β represents the mean value of citizens' preferences across the population, and η_i is the deviation from the mean of the preferences of citizen i . Although MIXL can accommodate any statistical distribution for β_i , under the N-MIXL model specification, it is assumed that β_i follows a normal distribution with mean $\underline{\beta}$ and covariance matrix $\underline{\Sigma}_\beta$, that is, $\beta_i \sim f\left(\beta_i \mid \underline{\beta}, \underline{\Sigma}_\beta\right)$.¹⁰

Therefore, the unconditional choice probability of observing choice j^* by citizen i in choice situation t is the expected value of the conditional logit probability over the parameter values.¹¹ This is the integral over all possible values of β_i , weighted by the distribution of β_i :

$$E(P_{ij^*t}) = \int P_{ij^*t}(\beta_i) * f\left(\beta_i \mid \underline{\beta}, \underline{\Sigma}_\beta\right) d\beta_i \quad (7)$$

M3. Theoretically restricted mixed logit model (T-MIXL)

As shown in McFadden and Train (2000), if the mixing distribution is chosen appropriately, the MIXL random utility model can be approximated by this econometric specification. From a theoretical point of view, this result is very appealing and has been stated as the main argument for adopting a MIXL in DCE applications during the last decade or so. However, from an empirical point of view, it is implausible to test the potential alternative specifications of MIXL models. The rest of the modeling specifications that have been employed in the literature—and in particular those that we are using in this analysis—can be seen as specific cases of the MIXL model with different mixing distributions. One of these models is the T-MIXL. Unlike the N-MIXL in which all the parameters are assumed to follow a normal distribution, the T-MIXL model incorporates an additional restriction for the price parameter. Since for most DCE applications, economic theory suggests that it is very unlikely that rising prices will positively affect preferences, T-MIXL assumes that such parameters follow a constrained triangular distribution. This assumption ensures that only negative values of the parameters are considered (Greene et al. 2006). We also explored the use of alternative specifications for the price parameter (i.e. log-normal, specific utility forms,...). However, since the main results of this study were not sensitive to this choice specification, here we provide the most common one involving using a constrained triangular distribution.

¹⁰ It is worth noting that this specification is equivalent to assuming that η_i follows a multivariate normal distribution with zero mean and a diagonal covariance matrix.

¹¹ Since this integral does not have a closed-form solution, McFadden and Train (2000) suggested to estimate the model using simulated maximum likelihood methods or Markov Chain Monte Carlo simulation techniques.

M4. Scaled multinomial logit models (SMNL)

In all the above econometric specifications of the RUM, the scale parameter σ is normalised to one, in order to ensure parameter identification. However, several commentators have claimed that choice data are likely to present heterogeneity in the scale parameter in ways that are not explicitly captured by adopting N-MIXL or T-MIXL models (Swait and Louviere 1993). Fiebig et al. (2010) proposed an alternative model specification that allows researchers to account for heterogeneity across respondents in the random component of utility. This model has been termed elsewhere as the scaled multinomial logit (S-MNL) model (Keane and Wasi 2013; Hensher and Greene 2010; Hensher et al. 2011).

The main characteristic of the S-MNL model is that the error variance is allowed to be heterogeneous in the population. Thus, the conditional indirect utility that citizen i derives from program j in choice situation t (U_{ijt}) is given by:

$$U_{ijt} = [\beta\sigma_i]' X_{ijt} + \varepsilon_{ijt} \quad (8)$$

where the parameters σ_i collect the specific standard deviation of the error term for each citizen i , capturing potential scale heterogeneity. In order to guarantee that individual scaling factors are strictly positive, we use an exponential transformation as in Fiebig et al. (2010), that is,

$$\sigma_i = \exp[E(\sigma_i) + \tau w_i] \quad (9)$$

where w_i follows a standard normal distribution and τ is a parameter accounting for the unobserved scale heterogeneity. Since t is directly and positively related with the existence of scale heterogeneity in the sample, its interpretation is that the higher the parameter, the higher is the likelihood that there is scale heterogeneity in the data set. Finally, in order to ensure identification, the error term is normalised such as $E(\sigma_i) = 1$; and therefore, $\sigma = -\tau^2/2$.

M5. Generalized mixed logit models (G-MNL)

The generalized mixed logit (G-MNL) is a flexible specification that allows researchers to accommodate individual scale as well as individual preference heterogeneity. It was firstly proposed by Fiebig et al. (2010). The G-MNL model specification nests N-MIXL and S-MNL by accounting for unobserved heterogeneity in both the systematic and the random components of the conditional indirect utility function. In the G-MNL model, utility U_{ijt} is defined by:

$$U_{ijt} = [\sigma_i\beta + \gamma\eta_i + (1 - \gamma)\sigma_i\eta_i]' X_{ijt} + \varepsilon_{ijt} \quad (10)$$

where σ_i is the individual-specific standard deviation of the error term capturing scale heterogeneity; η_i is individual-specific deviations from the mean, capturing individual heterogeneity in preferences; and γ is a parameter between zero and one,

that can capture how the variance of the individual preference heterogeneity varies with scale.

Finally, in order to estimate G-MNL, we need both to define a specification of the statistical distribution of σ_i and to include some restrictions to ensure parameter identification. Here, we follow the recommendations of Keane and Wasi (2013) and use a log-normal distribution that guarantees positive values for the scale parameter, that is, $\log(\sigma_i) \sim N(\beta, \tau^2)$. Note that parameters τ and β cannot be jointly identified. The strategy proposed in Keane and Wasi (2013) is to estimate τ and β , and then to calibrate accordingly. By doing this, β can be interpreted as a vector collecting mean preferences for each attribute in the choice set.

M6. The latent class model (LC)

Another popular model specification for analysing choice data is the latent class model (LC). In this model, preference heterogeneity is accounted for by a discrete distribution over unobservable endogenous (latent) classes of respondents (Boxall and Adamowicz 2002; Wedel and Kamakura 2000). Preferences are assumed to be homogeneous within each class but are allowed to differ across classes. The population is thus represented as consisting of a finite number of segments or classes (S). Respondents are allocated to segments simultaneously with the analysis of choices. The number of segments is often unknown before the experiment, so it is endogenously determined by the data, whilst membership of a segment depends probabilistically on the respondent’s observable socio-economic or attitudinal and behavioural characteristics. The most commonly employed criteria to decide the number of classes (S) are the Bayesian information criterion (BIC) or the Aikake information criterion (AIC).

In the random utility framework for the LC, the utility a citizen i who belongs to segment s derives from program j at moment t is given by,

$$U_{ijt|s} = \beta'_s X_{ijt} + \epsilon_{ijt|s} \tag{11}$$

where β'_s is the segment-specific vector of coefficients, X_{ijt} is the vector of attributes associated with each program and $\epsilon_{ijt|s}$ is the random component of utility for each segment. Since the vectors of coefficients differ between segments, preference heterogeneity across segments is captured. Under the assumption of independently and identically distributed (iid) error terms that follow a Type 1 extreme value distribution, the probability that program j^* is selected by a citizen i belonging to segment s is given by:

$$P_{ij^*t|s} = \frac{\exp(\beta'_s X_{ij^*t|s})}{\sum_{j=1}^k \exp(\beta'_s X_{ijt|s})}; \forall j \neq j^* \tag{12}$$

Membership of a specific segment is determined by a likelihood function M that classifies respondents to one of the segments with probability P_{is} . The membership

likelihood function is given by $M_{is} = a_s Z_i + \xi_{is}$; where Z_i is a vector of socio-economic and other observed characteristics of the respondent and ξ_{is} represents the error term. Assuming that this error term is also iid and follows a type 1 extreme value distribution, the probability that a citizen i belongs to segment s^* is:

$$P_{is^*} = \frac{\exp(a'_{s^*} Z_i)}{\sum_{s=1}^S \exp(a'_s Z_i)}; \forall s \neq s^* \quad (13)$$

The joint probability that citizen i chooses program j^* at the moment t is given by

$$P_{ij^*t} = \sum_s \left[P_{ij^*t|s} \right] * [P_{is}] = \sum_s \left[\frac{\exp(\beta'_s X_{ij^*t|s})}{\sum_{j=1}^k \exp(\beta'_s X_{ijt|s})} \right] * \left[\frac{\exp(a'_{s^*} Z_i)}{\sum_{s=1}^S \exp(a'_s Z_i)} \right] \quad (14)$$

M7. The mixture-of-normals mixed multinomial logit model (MM-MNL)

The use of mixture-of-multivariate-normals as an alternative flexible distribution is present in the literature. In Geweke and Keane (2001), the authors develop the mixture-of-normals probit model. This model has been applied by Araña and León (2005) in environmental valuation using dichotomous choice contingent valuation data (DCCV). In the context of DCE, a similar model has been applied to data sets from several choice experiments in different contexts (Brey and Walker 2011; Keane and Wasi 2013). The original application of the mixture-of-normals model to DCE data was probably Burda et al. (2008). In this paper, the authors specified a subset of coefficients in a MIXL model to follow mixture-of-normal distributions, whilst some others followed a simple normal distribution.

As Keane and Wasi (2013) point out, the MM-MNL (or “mixed-mixed logit”) model essentially nests the MIXL with LC models, with the aim of minimising the disadvantages of each. In fact, specifying the mixing distribution of MIXL to be mixture-of-normals is equivalent to extending LC models to incorporate unobserved heterogeneity within class. Thus, the utility of citizen i in period t conditional on choice of program j is specified as:

$$U_{ijt} = \beta'_i X_{ijt} + \varepsilon_{ijt}. \quad (15)$$

where β_i follows a $MVN\left(\beta_s, \underline{\sum \beta_s}\right)$ with probability P_{is} . $\sum_{s=1}^S P_{is} = 1$; and $P_{is} \geq 0, \forall s$; $s = 1, 2, \dots, S$. Note that when $S = 1$, MM-MNL boils down to the N-MIXL model presented above. On the other hand, when $\underline{\sum \beta_s} = 0$, MM-MNL becomes the LC model. Therefore, the choice probabilities are given by the following expression:

$$P_{ijt} = \sum_{s=1}^S P_{is} * \left\{ \int \left[\prod_t \left(\frac{\exp(\beta'_{is} X_{ijt|s})}{\sum_{j=1}^k \exp(\beta'_{is} X_{ijt|s})} \right) * f(\beta_{is} | \bullet) d\beta_{is} \right] \right\} \quad (16)$$

Funding Open Access funding provided by Vienna University of Economics and Business (WU).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adamowicz, W., Boxall, P., Williams, M., & Louviere, J. (1998). Stated preference approaches for measuring passive use values: Choice experiments and contingent valuation. *American Journal of Agricultural Economics*, 80(1), 64–75.
- Adamowicz, W. L., Glenk, K. & Meyerhoff, J. (2014). Choice modelling research in environmental and resource economics. In *Handbook of choice modelling*. Edward Elgar Publishing.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Akaike, H. (1987). Factor analysis and aic. In *Selected papers of Hirotugu Akaike* (pp. 371–386). Springer.
- Alberini, A., & Longo, A. (2006). Combining the travel cost and contingent behavior methods to value cultural heritage sites: Evidence from Armenia. *Journal of Cultural Economics*, 30(4), 287–304.
- Angelini, F., & Castellani, M. (2019). Cultural and economic value: a critical review. *Journal of Cultural Economics*, 43(2), 173–188.
- Apostolakis, A., & Jaffry, S. (2005). A choice modeling application for Greek heritage attractions'. *Journal of Travel Research*, 43(3), 309–318.
- Apostolakis, A., & Jaffry, S. (2005). Stated preferences for two Cretan heritage attractions'. *Annals of Tourism Research*, 32(4), 985–1005.
- Apostolakis, A., & Jaffry, S. (2007). The effect of cultural capital on the probability to visit cultural heritage attractions. *International Journal of Tourism Policy*, 1(1), 17–32.
- Araña, J. E., & León, C. J. (2005). Flexible mixture distribution modeling of dichotomous choice contingent valuation with heterogeneity. *Journal of Environmental Economics and Management*, 50(1), 170–188.
- Avrami, E., Mason, R., & de la Torre, M. (2000). *Values and heritage conservation: Research report*. Los Angeles: The Getty Conservation Institute.
- Balcombe, K., Bailey, A., Chalak, A., & Fraser, I. (2008). Modifying willingness to pay estimates where respondents MIS-report their preferences. *Applied Economics Letters*, 15(5), 327–330.
- Balcombe, K., Fraser, I., & Harris, L. (2009). Consumer willingness to pay for in-flight service and comfort levels: A choice experiment. *Journal of Air Transport Management*, 15(5), 221–226.
- Balcombe, K., Fraser, I., Lowe, B., & Souza-Monteiro, D. (2016). Attribute non-attendance and attribute importance ranking responses within discrete choice experiments. *American Journal of Agricultural Economics*, 98, 54–73.
- Balcombe, K., Fraser, I., & McSorley, E. (2015). Visual attention and attribute attendance in multi-attribute choice experiments. *Journal of Applied Econometrics*, 30(3), 447–467.

- Ben-Akiva, M. E., Lerman, S. R., & Lerman, S. R. (1985). *Discrete choice analysis: Theory and application to travel demand* (Vol. 9). Cambridge: MIT Press.
- Boxall, P. C., & Adamowicz, W. L. (2002). Understanding heterogeneous preferences in random utility models: A latent class approach. *Environmental and Resource Economics*, 23(4), 421–446.
- Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52(3), 345–370.
- Brey, R., & Walker, J. L. (2011). Latent temporal preferences: An application to airline travel. *Transportation Research Part A: Policy and Practice*, 45(9), 880–895.
- Burda, M., Harding, M., & Hausman, J. (2008). A bayesian mixed logit-probit model for multinomial choice. *Journal of Econometrics*, 147(2), 232–246.
- Campbell, D., Hutchinson, W. G., & Scarpa, R. (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.
- Caputo, V., Scarpa, R., & Nayga, R. M., Jr. (2017). Cue versus independent food attributes: The effect of adding attributes in choice experiments. *European Review of Agricultural Economics*, 44(2), 211–230.
- Chen, C.-F., & Chen, P.-C. (2012). Exploring tourists' stated preferences for heritage tourism services-the case of Tainan City, Taiwan. *Tourism Economics*, 18(2), 457–464.
- Choi, A. S., Ritchie, B. W., Papandrea, F., & Bennett, J. (2010). Economic valuation of cultural heritage sites: A choice modeling approach. *Tourism Management*, 31(2), 213–220.
- Clark, K., & Maer, G. (2008). The cultural value of heritage: Evidence from the heritage lottery fund. *Cultural Trends*, 17(1), 23–56.
- Crossick, G., & Kaszynska, P. (2016). Understanding the value of arts and culture: The arts and humanities research council project
- Czajkowski, M., Giergiczy, M., & Greene, W. H. (2014). Learning and fatigue effects revisited: Investigating the effects of accounting for unobservable preference and scale heterogeneity. *Land Economics*, 90(2), 324–351.
- Deodhar, V., et al. (2004). *Does the housing market value heritage?: Some empirical evidence*, Research Papers 0403. Department of Economics: Macquarie University.
- Fiebig, D. G., Keane, M. P., Louviere, J., & Wasi, N. (2010). The generalized multinomial logit model: Accounting for scale and coefficient heterogeneity. *Marketing Science*, 29(3), 393–421.
- Geweke, J., & Keane, M. (2001). Computationally intensive methods for integration in econometrics. In *Handbook of econometrics* (Vol. 5, pp. 3463–3568). Elsevier.
- Greene, W. H., & Hensher, D. A. (2010). *Modeling ordered choices: A primer*. Cambridge: Cambridge University Press.
- Greene, W. H., Hensher, D. A., & Rose, J. (2006). Accounting for heterogeneity in the variance of unobserved effects in mixed logit models. *Transportation Research Part B: Methodological*, 40(1), 75–92.
- Grisolía, J. M., & Willis, K. (2011). An evening at the theatre: Using choice experiments to model preferences for theatres and theatrical productions. *Applied Economics*, 43(27), 3987–3998.
- Grisolía, J. M., & Willis, K. G. (2012). A latent class model of theatre demand. *Journal of Cultural Economics*, 36(2), 113–139.
- Hanley, N., Wright, R. E., & Adamowicz, V. (1998). Using choice experiments to value the environment. *Environmental and Resource Economics*, 11(3–4), 413–428.
- Hensher, D. A. (1994). Stated preference analysis of travel choices: The state of practice. *Transportation*, 21(2), 107–133.
- Hensher, D. A., Beck, M. J., & Rose, J. M. (2011). Accounting for preference and scale heterogeneity in establishing whether it matters who is interviewed to reveal household automobile purchase preferences. *Environmental and Resource Economics*, 49(1), 1–22.
- Hensher, D. A., & Greene, W. H. (2010). Non-attendance and dual processing of common-metric attributes in choice analysis: A latent class specification. *Empirical economics*, 39(2), 413–426.
- Hutter, M., & Throsby, D. (2008). *Beyond price: Value in culture, economics, and the arts*. Cambridge: Cambridge University Press.
- Jaffry, S., & Apostolakis, A. (2011). Evaluating individual preferences for the British museum. *Journal of Cultural Economics*, 35(1), 49–75.
- Johnston, R. J., Boyle, K. J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T. A., et al. (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, 4(2), 319–405.

- Keane, M., & Wasi, N. (2013). Comparing alternative models of heterogeneity in consumer choice behavior. *Journal of Applied Econometrics*, 28(6), 1018–1045.
- Kragt, M. E. (2013). Stated and inferred attribute attendance models: A comparison with environmental choice experiments. *Journal of Agricultural Economics*, 64(3), 719–736.
- Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, 74(2), 132–157.
- Lazrak, F. (2009). *Cultural Heritage: Hedonic Prices for Non-market Values*. Serie research memoranda: Vrije Universiteit, Faculty of Economics and Business Administration.
- Liebe, U., Meyerhoff, J., & Hartje, V. (2012). Test-retest reliability of choice experiments in environmental valuation. *Environmental and Resource Economics*, 53(3), 389–407.
- Lourenço-Gomes, L., Pinto, L. M. C., & Rebelo, J. F. (2014). Visitors' preferences for preserving the attributes of a world heritage site. *Journal of Cultural Heritage*, 15(1), 64–67.
- Louviere, J. J., & Hensher, D. A. (1983). Using discrete choice models with experimental design data to forecast consumer demand for a unique cultural event. *Journal of Consumer Research*, 10(3), 348–361.
- Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: Analysis and applications*. Cambridge: Cambridge University Press.
- Louviere, J. J. & Woodworth, G. (1983). Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. *Journal of Marketing Research*, 350–367.
- Maddison, D., & Foster, T. (2003). Valuing congestion costs in the British museum. *Oxford Economic Papers*, 55(1), 173–190.
- Mas, A., & Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, 107(12), 3722–59.
- Mason, R. (2005). Museums, galleries and heritage. In *Heritage, museums and galleries: An introductory reader* (pp. 200–214). Routledge.
- Mason, R. (2008). Be interested and beware: Joining economic valuation and heritage conservation. *International Journal of Heritage Studies*, 14(4), 303–318.
- Mazzanti, M. (2002). Cultural heritage as multi-dimensional, multi-value and multi-attribute economic good: Toward a new framework for economic analysis and valuation. *The Journal of Socio-Economics*, 31(5), 529–558.
- Mazzanti, M. (2003). Valuing cultural heritage in a multi-attribute framework microeconomic perspectives and policy implications. *The Journal of Socio-Economics*, 32(5), 549–569.
- McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, 3(4), 303–328.
- McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of applied Econometrics*, 15(5), 447–470.
- Miller, S., Tait, P., & Saunders, C. (2015). Estimating indigenous cultural values of freshwater: A choice experiment approach to māori values in new Zealand. *Ecological Economics*, 118, 207–214.
- Morey, E. R., Rossmann, K. G., Chestnut, L. G., Ragland, S., Navrud, S., & Ready, R. (2002). Valuing reduced acid deposition injuries to cultural resources: Marble monuments in Washington, DC. *Valuing cultural heritage: Applying environmental valuation techniques to historic buildings, monuments and artifacts* (pp. 159–183). Cheltenham: Edward Elgar.
- Morey, E., & Rossmann, K. G. (2003). Using stated-preference questions to investigate variations in willingness to pay for preserving marble monuments: Classic heterogeneity, random parameters, and mixture models. *Journal of Cultural Economics*, 27(3–4), 215–229.
- Moro, M., Mayor, K., Lyons, S., & Tol, R. S. (2013). Does the housing market reflect cultural heritage? A case study of greater Dublin. *Environment and Planning A*, 45(12), 2884–2903.
- Peacock, A. (1994). The design and operation of public funding of the arts: An economist's view. In *Cultural economics and cultural policies* (pp. 167–184). Springer.
- Provins, A., Pearce, D., Ozdemiroglu, E., Mourato, S., & Morse-Jones, S. (2008). Valuation of the historic environment: The scope for using economic valuation evidence in the appraisal of heritage-related projects. *Progress in Planning*, 69(4), 131–175.
- Rigby, D., Burton, M., & Pluske, J. (2016). Preference stability and choice consistency in discrete choice experiments. *Environmental and Resource Economics*, 65(2), 441–461.
- Rizzo, I., & Throsby, D. (2006). Cultural heritage: Economic analysis and public policy. *Handbook of the Economics of Art and Culture*, 1, 983–1016.
- Rolfe, J., Bennett, J., & Louviere, J. (2000). Choice modelling and its potential application to tropical rainforest preservation. *Ecological Economics*, 35(2), 289–302.

- Rolfe, J., & Windle, J. (2003). Valuing the protection of aboriginal cultural heritage sites. *Economic Record*, 79(Special Issue), S85–S95.
- Scarpa, R., Campbell, D., & Hutchinson, W. G. (2007). Benefit estimates for landscape improvements: Sequential Bayesian design and respondents' rationality in a choice experiment. *Land Economics*, 83(4), 617–634.
- Scarpa, R., Gilbride, T. J., Campbell, D., & Hensher, D. A. (2009). Modelling attribute non-attendance in choice experiments for rural landscape valuation. *European Review of Agricultural Economics*, 36(2), 151–174.
- Scarpa, R., Thiene, M., & Train, K. (2008). Utility in willingness to pay space: A tool to address confounding random scale effects in destination choice to the alps. *American Journal of Agricultural Economics*, 90(4), 994–1010.
- Schaafsma, M., Brouwer, R., Liekens, I., & De Nocker, L. (2014). Temporal stability of preferences and willingness to pay for natural areas in choice experiments: A test-retest. *Resource and Energy Economics*, 38, 243–260.
- Schwarz, G., et al. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464.
- Strazera, E., Cherchi, E., & Ferrini, S. (2010). Assessment of regeneration projects in urban areas of environmental interest: A stated choice approach to estimate use and quasi-option values. *Environment and Planning A*, 42(2), 452–468.
- Swait, J., & Louviere, J. (1993). The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research*, 30(3), 305–314.
- Thiene, M., & Scarpa, R. (2009). Deriving and testing efficient estimates of WTP distributions in destination choice models. *Environmental and Resource Economics*, 44(3), 379.
- Throsby, D. (1999). Cultural capital. *Journal of Cultural Economics*, 23(1–2), 3–12.
- Throsby, D. (2001). *Economics and culture*. Cambridge: Cambridge University Press.
- Throsby, D. (2013). Assessment of value in heritage regulation. In *Handbook on the economics of cultural heritage* (p. 456). Edward Elgar Publishing.
- Throsby, D., & Zednik, A. (2014). The economic and cultural value of paintings: Some empirical evidence. In *Handbook of the economics of art and culture* (Vol. 2, pp. 81–99). Elsevier.
- Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34(4), 273.
- Train, K., & Weeks, M. (2005). Discrete choice models in preference space and willingness-to-pay space. In *Applications of simulation methods in environmental and resource economics* (pp. 1–16). Springer.
- Tuan, T. H., & Navrud, S. (2007). Valuing cultural heritage in developing countries: Comparing and pooling contingent valuation and choice modelling estimates. *Environmental and Resource Economics*, 38(1), 51–69.
- Tuan, T. H., & Navrud, S. (2008). Capturing the benefits of preserving cultural heritage. *Journal of Cultural Heritage*, 9(3), 326–337.
- Walker, M., & Marquis-Kyle, P. (2004). *The illustrated Burra Charter: Good practice for heritage places*. Australia Icomos.
- Wedel, M., & Kamakura, W. A. (2000). Mixture regression models. In *Market segmentation* (pp. 101–124). Springer.
- Willis, K., & Snowball, J. (2009). Investigating how the attributes of live theatre productions influence consumption choices using conjoint analysis: The example of the national arts festival, South Africa. *Journal of Cultural Economics*, 33(3), 167–183.
- Yao, R. T., Scarpa, R., Rose, J. M., & Turner, J. A. (2015). Experimental design criteria and their behavioural efficiency: An evaluation in the field. *Environmental and Resource Economics*, 62(3), 433–455.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.