The Economic Value of Culture: A Benefit Transfer Study

Report to the Department for Digital, Culture, Media & Sport

October 2018
Final report

Fujiwara, D., Bakhshi, H., Mourato, S., Lawton, R., Hotopp, U., Lagarde, A., Davies, J.

Contents

Acknowledgements: thanks go to Paul Crawford, David Fleming, Paul Kirkman, Clare Pillman, Harman Sagger, Xa Sturgis and Iain Watson for their assistance and to DCMS for their support of the project.

1 Introduction

1.1 Background

The main objective of this project is to understand the economic value that specific cultural institutions in England provide to individual visitors and non-visitors in the general population. This necessitates estimating both their use and non-use value. In Bakhshi et. al., (2015) we demonstrated in the case of two premier cultural institutions how, when primary data can be collected, economic valuation techniques can be used for this purpose. The aim of the present study is to test whether economic values for cultural sites can be transferred to similar sites in England.

Non-market values refer to the benefits provided by goods or services which are not traded in the market, i.e. are available to the public free of charge (and therefore have no market price) (Arrow and Solow 1993; Bateman et al. 2002). An understanding of the value of non-market goods and services to the public is important for public investment decisions, policy appraisal and evaluation in existing and future museums (HM Treasury 2011). Failing to adequately value these benefits risks an under-appreciation of the social value of cultural investments (Crossick and Kaszynska 2016).

The Department for Digital, Culture, Media and Sport (DCMS) (and their Culture and Sport Evidence (CASE) partners) require robust valuations of cultural goods and services in order to demonstrate in quantitative terms the value that they generate for users (use value) and for non-users (non-use value) in a manner consistent with best practice valuation methodology within government as set out in the HM Treasury Green Book.

To this end, we:

a) estimate the economic benefits associated with the services offered by four museums, and

b) test the transferability of the estimated values between institutions to investigate the potential of applying the values to additional museums without the need to conduct fresh valuation studies.

To estimate the economic benefits we apply the contingent valuation (CV) methodology employed in our earlier study for the DCMS and the Arts and Humanities Research Council (AHRC) Cultural Value Project (Bakhshi et al. 2015). CV is an established stated preference approach recognised by the HM Treasury Green Book. CV surveys elicit from people who use the good or service how much value they place in accessing the service, and those in the general population who do not use it are asked for the value they place on its continued existence (Bateman et al. 2002).

To assess transferability, we apply a technique known as value transfer or benefit transfer (BT) (Brouwer 2000). BT is the exercise of transposing 'primary' research findings from one site to another. It offers policy-relevant values in a fast and cost-effective way, by transferring information from the studied sites to others.

The study only includes the benefits to residents in England, excluding foreign visitors (in compliance with the HMT Green Book). While in some exceptional cases the benefits to foreign visitors can be included, this is not suitable for the present study, which is designed to be used to inform the allocation of funds raised by domestic taxes to cultural institutions.[1]

1.2 Research stages

The full study consists of two phases:

i. Scoping work to determine values associated with four institutions of a single cultural category (in our case museums), as well as to test the data collection methods, and probe the survey instrument design.
ii. Conducting stated preference valuation studies for the four cultural institutions. A sufficient number of survey responses are collected on a minimum number of institutions to allow for the testing of benefit transfer.

1.3 Site selection

We select four museums in England from a list of 22 candidate institutions. Sites are selected based on the following criteria:

- **The good valued:** What is the good/service that is being valued? For example, the value for visitors of being able to access a site/institution; or the value for non-users in the general population related to maintenance of collections for future generations.
- **User population:** Does a clear and identifiable visitor population exist for surveying use values?
- **The provision of the good:** Is the cultural institution/site excludable, i.e. can a hypothetical scenario be designed in which access to the site can be withdrawn if certain conditions are not met (e.g. can access be restricted via an entry fee)?
- **Homogeneity across time:** Have the collections and services offered by the institution changed significantly over the past three years (e.g. through large-scale refurbishment), which would create substantive differences in the experiences of visitors at different points in the visitor sample timeframe?
- **Homogeneity across institution:** Are the sites in each category a coherent group to permit benefit transfer? Is there free entry to all the sites? Is the subject area of the collections comparable (e.g. historical museums)? Are they commensurate in importance (e.g. cultural institutions of local, national, or international significance)?
- **Payment mechanism:** Is it possible to design convincing hypothetical scenarios that elicit:
  - **Use value**: Can, for example, access to the institution be made contingent on a hypothetical visitor entry fee?
  - **Non-use value:** Can the maintenance of collections and presentation standards be made conditional on receiving donations?

[1] Note that this study does not include an assessment of the wider economic benefits of cultural institutions such as increased business for local shops caused by the visitors attracted to a city.

Compulsory mechanisms like entry fees or taxes are generally preferable to voluntary mechanisms like donations because they are in principle incentive compatible within the axioms of welfare economics, i.e. they incentivise truth-telling (Bateman et al. 2002). However, for non-use values the mechanism of an entrance fee becomes problematic. A scenario in which we ask respondents who have not visited the museum and may not visit in the future to express their willingness to pay via an entry fee is not relevant. On the other hand, a tax to maintain free access to the museum would be more appropriate if the question was framed in terms of a national policy, i.e. whether the public would be willing to pay to maintain free access to all museums and not just a single museum. But in the case where we are interested in a single institution, a voluntary payment mechanism, such as a donation to support the work of the museum in maintaining its collections, is arguably the most credible means of eliciting values for non-users (see e.g. Bedate et al. 2009; Pung et al. 2004).

- **National spread:** Is there an appropriate geographical spread of sites across England?

We provide a table summarising the site selection process for phase 1 of the project (Annex A1). Based on combinations of the criteria outlined above, a final set of four museums is selected: the Great North Museum in Newcastle, the Liverpool World Museum, the National Railway Museum in York, and the Ashmolean Museum in Oxford. All of these museums are free to enter, have collections of national relevance, are based in major population centres outside of London, and have not undergone major refurbishment in the past three years.

1.4 Valuation

CV is a commonly used stated preference method for valuing non-market goods, as recommended by the HM Treasury Green Book. At the heart of CV is the careful design of surveys asking respondents directly to report their maximum willingness to pay (WTP) (for positive outcomes or to avoid negative outcomes) or minimum willingness to accept (WTA) compensation (for negative outcomes or to forego positive outcomes). The CV methodology has developed over a number of decades, introducing a range of best practice techniques to improve the robustness and welfare consistency of the non-market values elicited (Bakhshi et al. 2015; Bateman et al. 2002; Arrow and Solow 1993). Notwithstanding the many challenges, the advantage of CV over other forms of non-market valuation, such as revealed preferences (which use existing market prices as a proxy for the non-market good), is its wide applicability and flexibility: it can estimate both use and non-use values, as well as being applicable to a very wide range of goods and services, including estimating values associated with future changes (Bateman et al. 2002).

1.4.1 Use and non-use values in the context of cultural institutions

We refer to the widely used total economic value (TEV) framework, when considering the economic value of cultural institutions. The TEV categorises values into two main categories, use values, and non-use values (Pearce, Atkinson and Mourato, 2006; Eftec 2005; Mourato and Mazzanti 2002):

- **Use values** are subdivided into **direct** and **indirect use**. *Direct use* benefits could include recreational, leisure, and entertainment activities, as well as education, inspiration and knowledge. *Indirect use* benefits could arise in the form of enhanced community image, sense of place, and social interaction. A so-called *option value* can also be attached to potential future use of the services that cultural institutions provide (Mourato and Mazzanti 2002).

- **Non-use values**: Users may also hold *non-use values* for museums that they have not visited. We can identify a primary categorisation of non-use values associated with cultural institutions. **Non-use** values can be described as: ***altruistic*** *values* – welfare increases from knowing that others living will benefit; ***bequest*** *values* – welfare increases associated with knowing that future generations will benefit; and ***existence*** *values* – associated with welfare enhancements from knowing that the cultural institution, its services and collections, exist even if an individual does not experience a use benefit now or in the future (for example, this may refer to a sense of pride associated with the existence of a regional museum).

Many of the multiple benefits listed here are by their nature bundled together. When asking individuals to consider the value of a visit to a cultural institution, for example, it is difficult to meaningfully disentangle the value attached to recreation, to education, to visual amenity, to inspiration, etc. Partial separate identification of some of the broader benefits categories (e.g. use and non-use) may however be possible, with careful sample selection and survey design, but even in these cases use and non-use values are commonly conflated (see e.g. Bakhshi et al. 2015).

For ease of understanding, in this study we use the following terminology for consistency:

Use value refers to the willingness to pay stated by those who have visited the cultural institution within a designated time period. While these are expected to be *primarily* use values, we acknowledge that visitors may also hold non-use values for the preservation and maintenance of collections. Use value within this study refers exclusively to the WTP values held by museum visitors (users) for accessing the museum.

Non-use value refers to the willingness to pay stated by those who have not visited the cultural institution within a designated time period. While these are expected to be *primarily* non-use values, we acknowledge that non-visitors may hold elements of use value, such as the option value to visit the museum in the future or having used it online for research or recreational reasons. Non-use value within this study refers exclusively to the WTP values held by museum non-visitors (non-users) for the museum's on-going conservation and maintenance work.

1.5 Eliciting Willingness to Pay

Stated preference surveys require that respondents find the hypothetical scenario to be meaningful and realistic, and that it reflects as closely as possible a real-life payment decision. This requires a realistic payment vehicle: the means by which their stated value would go toward securing the continued presence of the good or service (Bateman et al. 2002). We use two payment vehicles in this study: one to elicit museum visitors' use value the other to elicit the non-use value of those who have not visited the museum in question.

Specifically, we elicit museum visitors' **use value** through a hypothetical entrance fee to secure continued access to the respective museum. And we elicit the **non-use value** of those who have not visited the museum in question through a hypothetical voluntary annual donation required to secure the museum's ongoing conservation and maintenance work ("to ensure that collections are adequately conserved and maintained, and presented in the best possible way").

In CV surveys, in order to elicit welfare-consistent values it is vital that the hypothetical scenario is believable, and that it justifies the introduction of the payment mechanism. The entry fee to a museum provides incentive compatibility, in that respondents would be unable to access the museum without payment, and is believable and credible, given that respondents are likely to have paid an entrance fee elsewhere.[2] A context of financial constraints in government spending arguably provide a credible explanation for the hypothetical introduction of entrance fees.

An annual donation was selected as the most appropriate payment vehicle for those who indicated that they had not visited one of the museums ('museum non-visitors'). Cheap talk scripts and/or oath scripts are used to incentivise truth-telling and certainty, and follow-ups to screen unreliable responses are introduced into the survey to overcome the well-documented problem of incentive incompatibility and free-riding behaviour in voluntary donations (Champ and Bishop 2001).[3]

The elicitation method is a payment card. Specifically, we present respondents with a card offering them a range of monetary amounts they can choose from. This provides respondents with a visual aid and helps remove starting point bias (Bateman et al. 2002; Maddison and Foster 2003; Maddison and Mourato 2001).

1.6 Literature review

1.6.1 Contingent valuation in the cultural sector

Empirical research eliciting economic values or benefits associated with access, preservation or restoration of cultural assets dates at least back to the 1980s when the first contingent valuation studies in the field were conducted, focusing on the arts, theatre, historical sites, museums, galleries, libraries and broadcasting (for a review, see Noonan 2003; Pearce and O'zdemiroglu 2002). Since then, many studies in the cultural sector have been conducted worldwide investigating a variety of benefits, both tangible and intangible. For a detailed literature review of these studies, see Bakhshi et al. (2015).

Bakhshi et al. (2015), as part of the AHRC Cultural Value Project, performed a large-scale empirical comparison of contingent and subjective wellbeing valuation in the context of the UK's cultural sector, eliciting visitor and general population willingness to pay for the use and non-use aspects of two cultural institutions: the Natural History Museum (NHM) in London and Tate Liverpool (TL) gallery through face-to-face visitor and online general

[2] A range of options exist, including national or local taxes, a tourist levy or voluntary donations. As discussed earlier, the alternative of a general tax mechanism is more commonly associated with the pool of funding for all museums at the regional or national level, and may therefore be confusing in a survey that elicits valuations at the level of individual institutions.

population surveys. The study also applied subjective wellbeing analysis, in terms of momentary wellbeing, testing for associations between activities performed in the past hour and levels of self-reported happiness and sense of purpose. Visitor use values were estimated as £6.65 on average for the NHM (as a hypothetical entry fee) and £10.83 for TL (as an annual donation to support the work inside the gallery). Average visitor non-use value to support the research and conservation work of the NHM was elicited as a voluntary top up donation (mean £2.78), while visitor non-use value of the work of TL in the wider community, elicited as a donation, averaged £8.00. The online survey captured non-use and option values for the general UK population (excluding Northern Ireland) as an annual donation. The survey design developed in Bakhshi et al. (2015) is used as the basis of the present survey design. This provides the advantage that survey design and wording has been tested in the field prior to the present study.

1.6.2 Benefit transfer: Literature review and applicability

Benefit transfer is the exercise of transposing 'primary' research findings from one site to another. It offers a means to providing policy-useful values in a fast and cost-effective way, by taking the estimated average WTP values from study sites and applying them to the new policy site (unit estimate transfer), or transferring the information from the study site to the policy site regarding the relationship between WTP and a number of explanatory variables (function transfer) (Brouwer 2000; Eftec 2000).

BT has been widely reviewed and discussed in the academic literature since it was first formalised in the 1980s as a technique for transferring point estimates (mean values) to other sites without the need for primary research. A set of more complex methodologies emerged in the 1990s with techniques that improved the robustness and flexibility of transfer to sites with different observable characteristics, achieved with the transfer of WTP functions (Atkinson et al. 1992; Boyle and Bergstrom 1992; Desvousges et al. 1992; Loomis 1992a; Smith 1992; Walsh et al. 1992). A full review of benefit transfer as developed since the 1990s can be found in Johnston et al. (2010; 2015).

Desvousges, Johnson and Banzhaf (1998) developed the first set of tests for analysis of the validity of BT. Over the last decade there has been a growing consensus around methods for reducing error in BT, through Brouwer (2000), Bateman et al. (2011) Brouwer et al. (2015), and Johnston et al. (2015).

The key challenge for BT methods is to avoid errors that lead to improper inferences regarding welfare effects and thereby misguided policy decisions. These include errors transferred from the original primary studies (*measurement errors*) and errors generated by the transfer process itself (*generalisation errors*). The larger the set of study sites, the lower the risk of measurement error related to the possible selection of a single inaccurate or inappropriate source study.

There are three broad approaches to BT (Brouwer 2000). The first is based on a transfer of a known benefit to another site, and the second on the characteristics of the users and non-users of a site and how much these characteristics are associated with the valuation of cultural and heritage goods. Finally, a third approach known as meta-analytic value function transfer uses a value function estimated from multiple study results, together with information on parameter values for the policy site, to estimate policy site values. The value function therefore does not come from a single study but

from a collection of studies. This allows the value function to include greater variation in both site characteristics (e.g. socio-economic and physical attributes) and study characteristics (e.g. valuation method) that cannot be generated from a single primary valuation study (Johnston et al. 2015). However, meta-analytic benefit function transfer requires greater expertise on the part of the policy analyst and more extensive data. We do not expand on this method further as it is not applied in this report.

A growing consensus exists on the advantages and disadvantages of the unit value and function transfer approaches (Johnston et al. 2015):

Table 1.1 Advantages and disadvantages of unit and function benefit transfer approach (from Johnston et al. 2015)

Unit value transfer	**Adjusted unit/Benefit function transfer**
Advantages	**Advantages**
Involves little or no modelling.	Increased flexibility and capacity to adjust welfare measures.
Less sensitive to modelling assumptions.	
Disadvantages	**Disadvantages**
Unable to adjust welfare measures according to characteristics of the policy site.	Over-parameterisation of model can introduce measurement error.
Least accurate transfer method on average, although performs acceptably when policy and study sites (including population characteristics) are very similar.	Matching of characteristics of study and policy sites still required.

Boyle et al. (2010) outline two primary 'rules' in the benefit transfer literature.

First, transfer errors are reduced where study sites and policy sites are similar in terms of their physical sites and populations' characteristics. The transfer of benefits is expected to be more accurate when the researcher can control for as many factors as possible (Bergland et al. 2002). In particular, criteria for reliable transfers (i.e. low transfer errors – either *ex post*, once data on the policy site becomes available, or when assessing transfer errors *ex ant*e in a study which uses a proxy on which data is available for an unknown policy site) include: (i) using the same survey instrument across study sites; (ii) valuing the same type of policy change and sharing similar property rights; (iii) conducting surveys at the same point in time, and (iv) having samples of respondents with similar cultural and social characteristics.

Our proposed benefit transfers broadly meet these criteria: the same survey instrument will be used; the same policy change is valued; the surveys are administered at similar points in time, and the user and general populations are similar. However, we are limited by the amount of information we have on the actual visitors to the museums compared to the respondents to our survey who visited those museums.

Second, function transfers can in principle lead to more accurate transfer estimates because of the ability to adjust the estimates according to observable differences between the sites (Johnston et al. 2015).

However, although the function transfer approach has the potential to be more robust and to provide lower errors (Rosenberger and Loomis 2003), this is not always the case in practice. Previous meta-review studies (Kaul et al. 2013; Whitehead et al. 2015) have failed to find consistent evidence that function transfer outperforms unit transfer, while others have even found that unit value transfer outperforms benefit function value transfer (Ready and Navrud 2006; Bergland et al. 2002; Brouwer and Spaninks 1999). There are other examples from the literature where comparison of unit and function transfer approaches has seen the value function transfer to increase transfer errors (Kristofersson and Navrud 2007). We review the recommended tests for BT in Section 4.1.

There are important examples of BT being applied in policy in the United Kindom, the European Union (see Brouwer and Navrud 2015 for a review), and the United States (see Loomis 2015 for review). In the UK, Eftec has produced detailed BT guidance on the use of value transfer in policy and project appraisal for Defra (Eftec 2009).

BT is used in the valuation of health impacts through Value of Statistical Life (VSL) unit values established as part of the European Commission ExternE project.[4] The European Commission initiated a review of all valuation studies of transportation noise to establish unit values per decibel (dBA) for amenity loss due to traffic noise (Navrud 2002). Brink et al. (2011) applied unit values from a meta-analytic value function to estimate the total economic value of implementing the EU's network of nature protection areas (Natura 2000). More recently, the UK National Ecosystem Assessment developed a meta-analytic value function to evaluate the economic value of the identified flows of ecosystem services (Bateman et al. 2011). Internationally, the BT approach was applied in the recent OECD report on health costs of air pollution (OECD 2014) Finally, the Environmental Valuation Reference Inventory (EVRI)[5] is supported by a number of OECD governments to maintain a database of benefit estimates.

While BT is widely used in the environmental and health valuation fields,[6] applications of BT to the valuation of cultural assets are rare (Eftec 2005; Mourato et al. 2014).

The main reason is arguably that some types of immoveable (e.g. monuments, archaeological sites) and moveable cultural assets (e.g. paintings, sculptures) are considered unique in terms of their characteristics, such as historic context, or national and international significance. This uniqueness reduces the potential for transferring their values to other contexts (Provins et al. 2008). Nevertheless, BT arguably provides a useful alternative to conducting original valuation where the cultural good and the respective policy change are of similar nature and significance (e.g. non-iconic historic buildings with similar architecture, exterior and interior decorations; indoor collections sharing similar types of objects/collections) or where the range of services/benefits provided are similar.

[4] www.externe.info
[5] https://www.evri.ca
[6] Detailed reviews of BT can be found in (Ståle Navrud and Ready 2007) and (Desvousges et al. 1998).

In this study, we choose four museums that are homogenous with respect to a number of characteristics (see Section 1.3) and arguably provide similar types of services, thereby increasing the scope for transferability and for subsequent use in project appraisal.

There have been previous attempts to assess the scope of using valuation techniques in the cultural heritage sector (Eftec 2005). Eftec conducted contingent valuation studies of a number of built heritage sites[7] in the UK for the purposes of BT, aiming at building a bank of values that could be applied to similar heritage sites in the UK. Each case study set out a description of the heritage asset, a description of the action that is to be appraised (e.g. restoration work), and the hypothetical scenarios related to closure of sites to the public and deterioration of heritage assets and collections. Use and non-use values were elicited from visitors and the general public.

Mourato et al. (2014) performed a comprehensive analysis using BT methods. They considered the economic benefits associated with reducing climate change damage to built heritage interiors in Europe. The study included ten heritage sites (including historic houses, museums and churches) across five countries (the UK, Sweden, Germany, Romania and Italy), testing for transfer errors within categories of cultural site using three transfer tests: simple unit transfer; adjusted (income differential) unit transfer; and function transfer. Transfer tests were performed for both use and non-use values. The authors found that heritage conservation values could be successfully predicted via value transfer approaches.[8] This extensive study of different categories of cultural institution across a number of countries strongly informs the design of the current study.

In sum, there is broad agreement on the transfer methods available for BT and the statistical tests that should be applied to assess the reliability of transfer from a set of study sites to a policy site which has not been valued. BT has been applied extensively in the field of environmental valuation, as well as in the heritage sector, but to date has not been applied to cultural institutions like museums. In designing the current study, we were informed by best-practice CV survey design in the cultural sector (Bakhshi et al. 2015) and recent large-scale benefit transfer of heritage sites within the European Union (Mourato et al. 2014).

We can in principle transfer values across time, across sites and across populations. In this report, for use values across institutions and populations we test the unit value, the adjusted unit value and the benefit function transfer approaches. For non-use values across institutions and for the same population (i.e. we use the same sample of the national population when investigating non-use values), we illustrate the transfer procedure using the simple unit transfer method, as the WTP functions have weaker explanatory power in this case. With larger data sets and more case studies, where we can control for museum-specific variables, the benefit function transfer approach should also be performed for non-use values.

Moreover, there are also important considerations in relation to the additivity of non-use values for multiple sites in our study. Within the survey instrument we ask respondents to give their willingness to make an annual donation for one non-visited museum only.

[7] Denbigh Townscape; Kennet & Avon Canal; Battersea Park; Lincoln Cathedral; Sandal Castle
[8] Mourato et al. (2014)

We do not collect non-use values for any additional museums, nor do we elicit their WTP to support all museums in England. Therefore, while the non-use values elicited here give a representation of the value that non-users in the general population may hold for individual museums, great care should be taken in aggregating these non-use values across multiple individual level sites at the national level.

In the remainder of this report, Section 2 details the data collection process and our research methodology and sets out the steps we take for dealing with potential sources of bias introduced by our survey design. Section 3 presents the main CV findings, describes our robustness checks and assesses our success in mitigating against biases. Section 4 tests the transferability of the WTP estimates from the sites surveyed in this study to potential policy sites and makes recommendations for benefit transfer.

2 Data Collection and Research Methodology

This section sets out the methodology used in this study. It starts with the description of the survey instrument, explains the data and units of analysis drawn from the observations, the various steps in the analysis, and finally discusses the potential biases and order effects (and what we have done to try and mitigate against them). It includes a brief description of the methods used to assess the transferability of values (developed in full in Section 5).

2.1 Survey

One of the main innovations of this study is in the design of a single online survey instrument to collect responses from **visitors** and **non-visitors** for each of the four selected museums.[9] A single online survey was on balance considered the most cost-effective way to collect primary data for multiple museum sites, compared with individual online surveys for each institution, or face-to-face visitor survey at the sites. In addition, online surveys help avoid the bias which occurs in face-to-face surveys where respondents provide 'socially acceptable' responses that they think the interviewer wants to hear.

[9] We use the online-panel provider Toluna. Online survey panels do not provide a true random probability sample, but permit quotas to be set on a range of relevant attributes, such as gender, age, location and socio-economic group. The benefit of quota sampling is that quotas can be set to mimic population demographics to make the survey representative on those chosen characteristics, or they can be set so as to over-sample groups of particular interest. For the Toluna panel, sample selection is made randomly using the profile criteria specified by us. When doing this Toluna takes account of predicted response rates by target demographic and region to avoid over-contacting panellists and to ensure that they do not introduce a bias in their responses. Historical propensity to answer surveys is not used to select a sample. Respondents are randomly selected for surveys that they have a likelihood of qualifying for. In addition, in order to mitigate category overuse and other forms of awareness bias, Toluna can exclude any panellist from a client's survey by topic of survey recently taken, frequency of participation, or for tracking studies' participants from prior waves. Finally, Toluna has developed a sampling technique that can be used to further ensure that respondents are representative of the desired target population. This technique compares the incoming sample to a baseline and allows respondents into the survey as long as they match what is in the baseline sample.

2.1.1 Survey instrument

The survey is divided into four sections.

The **first** section contains background questions on respondents' membership of heritage, conservation or environmental organisations, their recent use of other museums and cultural sites (past twelve months), prioritisation questions on the allocation of government funding, and a set of questions on cultural engagement, developed from the DCMS's Taking Part survey. A set of statements about the value of culture and museums are presented on a five-point Likert scale (1, strongly disagree; 5, strongly agree). The responses to these questions are used to assess the theoretical validity of the willingness to pay values.

In the **second** section, we ask respondents about their previous visits to any of the four museums selected through site scoping (recall Section 1.3). This includes questions on which, if any, of the four museums they have previously visited. We then ask which of the selected museums they have visited in the past three years. A three year-period is selected to ensure sufficiently large sample sizes. The experience of an individual who visited one year or three years ago is assumed not to differ significantly (none of our four sites have undergone large-scale refurbishment in that period).[10] We include follow-up verification questions on when the visit took place (including an option to revise their previous answer if they had not in fact visited) and how often they visited each institution in their lifetime. These questions are designed to eliminate potential overstatement or incorrect responses due to recollection bias (whereby the respondent recollects having visited more often or more recently than they in fact have due to the focussing effect introduced through the survey (as identified in previous studies, e.g. Bakhshi et al. 2015)).

The **third** section includes the valuation (willingness to pay) questions. The number of WTP valuations questions a respondent is asked depends on their visit history in the last three years. For example, those who indicated that they have visited two of the museums in the past three years are asked two valuation questions related to those museums (eliciting museum visitor use WTP), as well as one WTP question related to a museum which they have not visited (eliciting non-use WTP for one non-visited museum). One of the museums they have not previously visited is randomly selected and the corresponding non-visitor questionnaire then displayed (unless they report to have visited all four museums within the past three years).

Respondents are presented with information about the museums they have been asked to value.[11] This information includes the services and visitor attractions the museum offers, and the work it does to conserve and maintain its collection. The same information is shown to museum visitors and non-visitors alike. Respondents are asked about their familiarity with this information on a five-point Likert scale (1, not at all familiar; 5, extremely familiar). The WTP scenarios, payment mechanisms, and elicitation methods

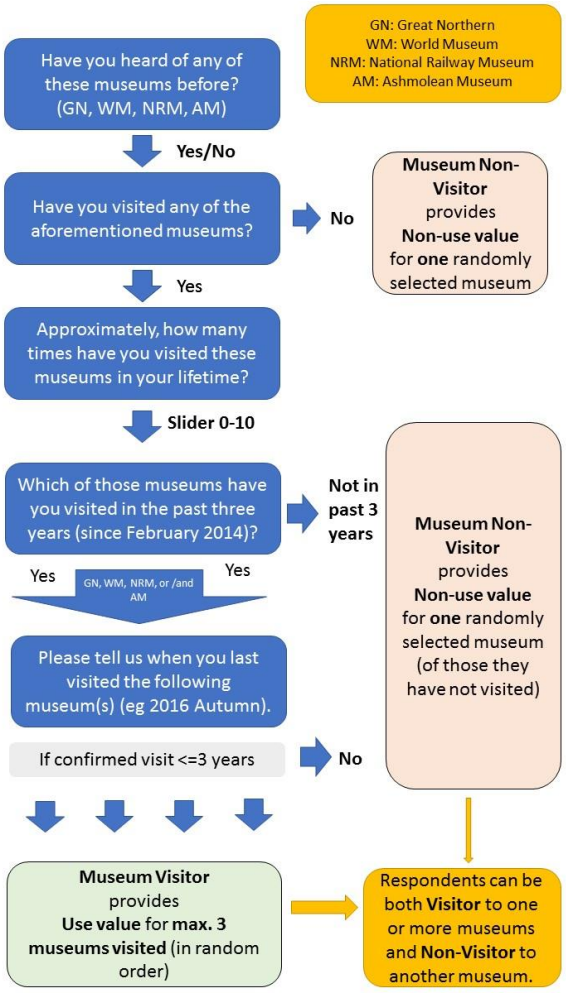
[10] The length of the recollection period (3 years) is used to allow a reasonable sample size for visitors. Mourato et al (2014) use the last 12 months for the reference period within which respondents have visited the site. However, there is no theoretical reason why users need only be defined as having visited within the past year. The key factor on deciding the acceptable length of time for defining users is the strength of their ability to recollect the goods and services offered by the institution being studied.

[11] Museum information is developed and signed off by representatives from each of the museums.

differed between the visitor and non-visitor questionnaires (as outlined in Sections 2.1.2.1 and 2.1.2.2). Figure 2.1 sets out the flow between Sections 2 and 3 in the survey. Within the survey design, respondents are always asked WTP of sites they have visited before the non-use WTP question.

The **fourth** section of the survey asks a set of standard socio-demographic questions, such as income, marital status etc, for use in validating the valuation estimates (Section 3.3.6 and 3.4.5) and in the BT analysis.

Figure 2.1 Identification of visitors and non-visitors within the survey instrument



2.1.2 Contingent valuation scenarios design

The survey instrument contains two CV questions: one to assess **Visitor Use Values** and a second to assess **Non-Visitor Non-Use Values**. We outline each set of CV questions in detail below.

2.1.2.1 Visitor questions – use values

Museum visitors are defined as those who have visited the institution in the past three years (as self-reported in Section 2 of the survey and excluding those who give inconsistent follow-up answers when asked for the date of visit).

Museum visitor WTP questions explore use values related to each of the museums the respondent indicates that they have visited in the last three years – for up to three museums in total.[12] Respondents are presented with a hypothetical scenario in which museums would need to find alternative sources of funding to support their activities, due to the current difficult financial situation. They are asked if in principle they would be willing to pay an entry fee to the museum under these circumstances. Those who answer positively (yes/maybe) are asked to state the maximum they would be willing to pay as an entry fee just for themselves to visit the museum (see Box 2.1). Museum visitor use WTP values are elicited using a payment card with values ranging from £0 to £75 with the smallest non-zero value being £0.25.

The same WTP question is asked for each of the museums that the respondent indicates they have visited, with information tailored to each museum. The order in which museums are presented is randomised to avoid possible order effects in the valuation (the depressing effect of each subsequent valuation on respondent WTP, for example, as their budget becomes constrained by previous payments or, in this case, as people become more aware of other museums they also value and might like to visit, see Section 2.4.3). This order is recorded for subsequent analysis of survey bias (Section 3.5). Respondents who have visited all four sites (n=19) are asked to value a maximum of three sites to keep the overall length of the survey acceptable.

[12] We restricted the use valuation to three institutions even if the respondent had been to all four museums in order to keep the survey length acceptable.

Box 2.1 Museum Visitor willingness to pay question (use value): entry fee

The difficult financial situation has meant that many museums in England have suffered cuts in funding while having to cope at the same time with increased numbers of visitors and associated increases in maintenance and operating costs.

In the unlikely event that funding ceases to be provided, museums would need to charge an entry fee to raise enough money to support their activities and secure their long-term future.

For the next set of questions, please imagine a situation where funding for museums has been cut and museum(s) start charging an entry fee. In this situation, please think about how much a visit to the INSTITUTION would be worth to you if anything.

Think of this as if you were going to the museum today and that this is the only museum you will visit today.

When answering this question, please ignore any other payment questions you might have answered in this survey and assume that this is the only payment you were asked for today. This is an entirely separate scenario for the INSTITUTION. The response should be strictly independent of any payment questions you may have already answered.

Would you be prepared to pay an entry fee to visit the INSTITUTION, even if this were only a small amount?

IF YES/MAYBE

If the INSTITUTION were no longer free, what is the maximum you would be willing to pay, as an entry fee, just for yourself, to visit the museum?

In answering this question, please focus solely on how much a visit to the INSTITUTION is worth to you, if anything, and how much you personally enjoyed the visit.

2.1.2.2 Museum non-visitor questions – non-use values

The museum non-visitor section of the survey is intended to explore mostly non-use and option values for museums that the respondent has not visited in the past three years (termed *non-use value* for clarity). One *museum non-visitor non-use WTP* question is randomly presented to each respondent (except in cases where a respondent had visited each of the four museums in the past three years). Randomisation of the non-visited museum site excludes any institutions the respondent indicated they have visited in past three years.

We provide respondents with information about the museum and its work maintaining and conserving the collections, using illustrative photographs, and ask how familiar they are with this information prior to the survey. Again, respondents are presented with a hypothetical scenario in which museums would need to find alternative sources of funding to support their activities due to the current difficult financial situation. In the case of museum non-visitors, these activities refer to the work it does to make sure the museum and its collections are adequately conserved and maintained and presented in the best possible way. Respondents are asked if they would be willing to pay in principle an annual donation to the museum under these circumstances. Those who answer positively (yes/maybe) are asked to state the maximum they would be willing to pay as an annual donation just for themselves to support this museum (see Box 2.2). Museum

non-visitor non-use WTP values are elicited using a payment card with values ranging from £0 to £75, with the smallest non-zero value being £0.25.

Box 2.2 Museum non-visitor willingness to pay question (non-use): annual donation

We now will provide you with some information about the funding of the INSTITUTION.

The difficult financial situation has meant that many museums in England have suffered cuts in funding while having to cope at the same time with increased numbers of visitors and associated increases in maintenance and operating costs.

Imagine if there were restrictions in funding because of the difficult financial situation, the INSTITUTION might have to consider alternative ways of funding the work it does to make sure the museum and its collections are adequately conserved and maintained, and presented in the best possible way.

For the next set of questions, please imagine a situation where funding for museums has been cut. In the unlikely event that funding ceased, please think about how much the INSTITUTION would be worth to you if anything.

When answering this question, please ignore any other payment questions you might have answered in this survey and assume that this is the only payment you were asked for today. This is an entirely separate scenario for the INSTITUTION. The response should be strictly independent of any payment questions you may have already answered.

Would you in principle be prepared to make an annual donation, even if only a small amount, to support the work the INSTITUTION does to make sure the museum and its collections are adequately conserved and maintained, and presented in the best possible way.

IF YES/MAYBE

Please think about how much the INSTITUTION is worth to you personally, if anything. What is the maximum you would be willing to pay, as a donation, per year, to support the work it does to make sure the museum and its collections are adequately conserved and maintained, and presented in the best possible way?

Both sets of valuation questions are followed by a certainty question where respondents are asked how certain they are about the amount they have stated on a sliding scale (0-100%): *How certain are you that you would really pay this amount if asked to enter the museum?/ would really donate this amount if asked?*

Respondents are also asked to select from a list of reasons for their willingness, or not, to pay. Again, these are used to assess the consistency of the responses. For instance, respondents who select that they gave their stated WTP because "*I don't believe that I would really have to pay"* are considered for exclusion from the set of values used to estimate mean WTP for that museum.

2.2 Data

This section sets out how we derive data from the respondents as they flow through the survey in terms of the museum-level units of analysis. The survey design provides **two study groups for WTP analysis: users (museum visitors) and non-users (museum non-visitors)**. Each of these broad groups have four sub-groups for each museum (as set

out in Table 2.1 and Figure 2.2, below). This provides eight different WTP valuation categories:

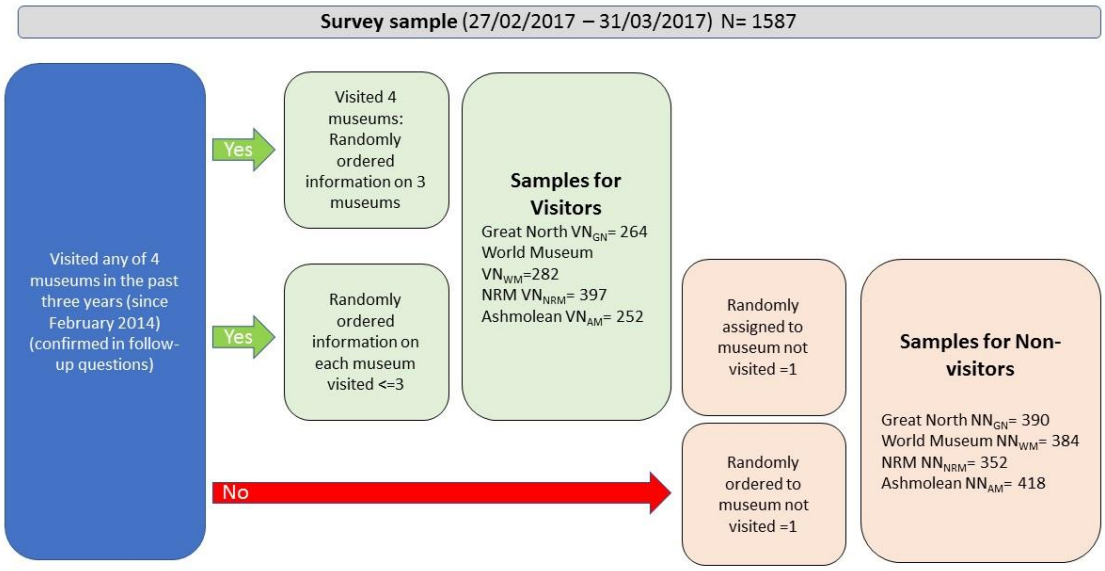
Table 2.1 Observation type and unit of analysis

Unit of analysis (museum)	Observation type and unit of analysis
Great North Museum, Newcastle (GN)	Visitor GN
	Non-visitor GN
World Museum, Liverpool (WM)	Visitor WM
	Non-visitor WM
National Railway Museum, York (NRM)	Visitor NRM
	Non-visitor NRM
Ashmolean Museum, Oxford (AM)	Visitor AM
	Non-visitor AM

The target sample for both visitors and non-visitors to each of the four museums is 250 per museum per user type. This sample includes people who have visited more than one museum, i.e. who provide WTP valuations for more than one museum. Therefore, the total number of WTP valuations is higher for some museums, meaning that completed museum-level WTP responses (visitor and non-visitors) across all four museums is larger (n=2,739) than the number of people who responded to the survey (n=1,587).

The allocation of respondents to different and multiple WTP questions means that *units of analysis* (number of observations) for the study are based on the WTP questions (use or non-use) for each of the museum institutions, rather than the total number of surveys completed (respondents). In other words, we estimate mean WTP at the *museum-level* (estimating the mean use/non-use WTP for all respondents who visited/non-visited each museum), rather than the individual-level (which would estimate mean WTP for each individual across all museums valued, which is not the objective of this study). This means that some individuals are present in multiple *museum-level* estimates of WTP (we outline the procedure for pooling together multiple museum-level WTP observations for the same individual in Section 2.2.1). Figure 2.2 outlines the flow from survey respondent to unit of analysis:

Figure 2.2 Survey flow: From respondent to observation and unit of analysis



Box 2.3 provides a summary of the study groups for analysis.

Box 2.3 Study groups

Users (museum visitors):

- Visitors to at least one of the museums in the past three years answer a **maximum of three use WTP valuation questions**. This sample includes the general population and booster sample. We apply museum visitor weights (tailored for each museum) to ensure representativeness with actual museum visitors.
- Within the survey design, respondents are always asked WTP of sites they have visited before the non-use question.

Non-users (museum non-visitors):

- Respondents who indicate that they have not visited a particular museum in the last three years.
- All respondents provide a non-use WTP (as an annual donation) for one museum that they have not previously visited (except in cases where individuals have visited all four institutions).
- We apply English general population weights (gender, age, region) from the Annual Population Survey (APS) to ensure representativeness with the national population.

2.2.1 Survey sample

Survey responses are screened to exclude people not resident in England and those under sixteen years old. We introduce loose quotas for gender, age and region to ensure that all

of these groups are represented in the survey. However, to ensure representativeness to the populations of relevance (museum visitors, or the general population of non-visitors), all responses are weighted. The weighting takes place in the following way. We obtain information on the breakdown by age and gender of the visitors to each of the four museums. Then, every respondent in our survey sample is assigned a weight according to the following formula:

Equation 1

$$w_i = \frac{p_i}{s_i}$$

Here, p_i is the share of the respondent's age-gender category in the target population, whereas s_i is the share of the same category in the survey sample. The target population for museum visitor WTP is the visitors of that particular museum aged 16 and above. The target population for non-visitor WTP is the whole population of England aged 16 and above, regardless of the museum(s) we are considering.[13] We drop n=3 respondents who indicate 'other' or 'rather not say' for the gender or age questions due to inability to calculate weights for them.

After getting these weights, we can calculate weighted means for WTP or any other characteristic *x* according to the formula:

Equation 2

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i w_i}{\sum_{i=1}^{N} w_i}$$

For the purposes of BT we combine all museum-level WTP values elicited from museum visitors and non-visitors into two pooled datasets, one for visitors, one for non-visitors.

There are several issues to consider when switching to pooled data analysis. Weights are set at the museum-level in terms of age and gender. When pooled data is used instead, the target population is a generic visitor to any museum of this type in England. Since demographic data on this category of people is not readily available, we use a second-best approximation by setting the target population for pooled visitor WTP to be the visitors of all four museums (aged 16 and above) joined together. For non-use values, the target population is the same for both the wide and the long data – the adult population of England, so there is no need to change the weights.

A further issue is that, by design, not every respondent provides a value for every museum. Moreover, the number of use WTP values provided varies across respondents. Out of the 1587 respondents, 691 (44%) provide no use values because they have not visited any of the museums. 633 (40%) provide one use value, 196 (12%) provide two use values and 66 (4%) provide three use values. A potential concern is that respondents who provide multiple use values are included multiple times in the analysis, thus making them over-represented in the pooled sample. We adjust for this over-representation by

[13] Office for National Statistics Annual Population Survey 2015-2016.

dividing their weights by the number of use values provided. For non-use values this over-representation of respondents is not an issue, as every respondent supplies only one non-use value. Only the 42 people who have visited all four museums provide no non-use values, and these are not included in the museum non-visitor study groups or pooled data at any point.

2.3 Analysis

2.3.1 Calculation of WTP

As described above, the visitor and non-visitor WTP questions provide two separate payment mechanisms (entry fee and annual donation). In all cases, use of a payment card elicitation mechanism means that respondents' stated values must be taken as a lower bound of their actual willingness to pay (Bateman et al. 2002) because the actual amount they are willing to pay will lie somewhere between the amount they choose and the next amount on the payment card.

We calculate mean and median WTP from the mid-point between the amount chosen on the card and the next amount up for use and non-use values within the eight study groups (museum visitors and non-visitors, recall Table 2.1). Following standard practice, all those who respond that they are not willing to pay in principle are coded as £0 bids (Bateman et al. 2002). Willingness to pay is calculated using museum-level weights (for museum visitor WTP) and national representative weights for non-visitors.

We report mean and median WTP, 95% confidence intervals, maximum values and the proportion of respondents giving a zero response, for each of the eight museum-level study groups. Following best practice in CV studies (Darling et al 2000), we develop BT testing using mean WTP only. The mean is relevant if the context of the valuation exercise is to aggregate values to the national level, because it represents an average WTP for the population which can be aggregated (by the population size) to derive the total WTP across the population.

We explore possible protest bids (invalid positive WTP values) by analysing the reasons given by respondents for being willing, or not, to pay. We remove from the sample individuals who give inconsistent reasons for their stated willingness to pay. For example, we classify as invalid responses those who state they don't 'believe [they] would have to pay' as an indicator of severe hypothetical bias. We assess the impact of this removal on the WTP value in sensitivity analysis. We find no significant difference in mean WTP with and without inconsistent responses across any museums.

2.3.2 Assessment of the validity of the WTP values

Following the calculation of mean and median WTP values we conduct validity tests to assess whether the values are valid within the existing theoretical context (including the wider academic literature, e.g. Bateman et al. 2002) and prior expectations around cultural engagement and past usage. For example, we expect individuals with higher income and those with an interest in culture to have on average higher WTP.

The following regression model is used as the base for the analysis of the validity of the WTP results (following Bakhshi et al. 2015):

Equation 3

$$WTP_i = \alpha + \beta_1 X_i + \varepsilon_i$$

where WTP_i is the amount the individual i states they are willing to pay, α is a constant and ε is the error term containing unobserved factors that determine willingness to pay. The X_i are the variables we use to control for the observed determinants of WTP (Bateman et al., 2002). These include those that are theoretically expected to affect WTP (such as income) as well as other factors that are known from the literature to have an effect e.g. positive attitudes towards museums and conservation (Bakhshi et al. 2015).

We estimate Equation 3 with socio-demographic variables, variables capturing experience of the museum (in the case of a use value), and attitudes, opinions, and proxy variables for cultural engagement (in all cases). We also include controls to capture the differences in survey mode, such as whether the respondent answers use and non-use WTP questions (in non-use WTP models) and the random order in which parts of the survey are presented. We estimate Equation 3 for use and non-use WTP measures respectively, as well as for pooled use (museum visitor) and non-use (non-visitor) samples, which provide greater sample size for the purposes of analysis.

We apply a number of tests for the validity[14] of our results, including testing the distribution of residuals for heteroskedasticity using robust standard errors and for normality using kernel density estimates (Annex A2). We highlight any results which may indicate that the WTP values are not valid within the sensitivity analysis.

2.3.3 Test for certainty

One subset of validity tests are tests for the certainty which respondents express when asked how certain they are that they would pay the stated fee to enter a museum (museum visitor use WTP) or the donation to support the museum (non-visitor non-use WTP) (Bedate et al. 2009). In Annex A3 we test for the association between certainty (measured as a percentage, where 100% is completely certain) and WTP. Specifically, we estimate the following model:

Equation 4

$$WTP_i = \alpha + \beta_1 Cert_i + \beta_2 X_i + \varepsilon_i$$

where WTP_i is the amount the individual i states they are willing to pay, $Cert_i$ is the individual's stated certainty to pay that value on a scale of 0-100%, X_i controls for standard socio-demographic determinants of WTP, α is the deterministic factor and ε_i is the error term containing unobserved factors that influence WTP.

We explore possible protest bids by analysing the reasons given by respondents for being willing, or not, to pay an entrance fee (museum visitor use WTP) or a donation to support the museum (non-visitor non-use WTP) (Annex A4). Since in the event the estimated number of protests is small, all responses are retained in the analysis.

[14] For more detail on validity tests see (Shadish et al. 2002)

2.4 Biases and Order Effects

In designing the CV scenarios (both museum visitor use WTP, and museum non-visitor non-use WTP) we take steps to attempt to deal with the known biases in CV and with the order effects specific to this study. The measures used in this study address biases which commonly occur in CV studies (Arrow and Solow 1993; Bateman et al. 2002; Carson 2012), with specific attention to order effects. We deal with each in turn.

2.4.1 Hypothetical bias

Respondents are randomly presented with entreaty scripts designed to reduce **hypothetical bias** and make the survey incentive compatible with standard welfare theory (Carlsson et al. 2013; Cummings and Taylor 1999a): they are asked if they promise to answer the WTP question truthfully (oath script, see Jacquemet et al. 2013), and are provided with cheap talk scripts asking them to be realistic, reminding them of the household budgetary constraints, and the existence of other cultural institutions that they may wish to spend their money on (Cummings and Taylor 1999a).[15] Respondents are also informed that "studies have shown that many people answering surveys such as this one, say they are willing to pay more than they would actually pay in reality" (Champ and Bishop 2001; Cummings and Taylor 1999b).

2.4.2 Starting point bias

Another important bias which can occur in stated preference studies is starting point bias. This bias means that the stated WTP may differ depending on the value at which the respondent starts to consider how much they would be willing to pay (Bateman et al. 2002). We apply best practice to the payment card we use for the valuation by starting at £0 and moving up the payment ladder in small steps (£0.25 and £0.5) up to a reasonable level of £75. The use of a payment card, compared to asking individual values, e.g. Would you be willing to pay more or less than £5?, removes the **starting point bias** as the respondents see a variety of values at the same time (Bateman et al. 2002; Maddison and Foster 2003; Maddison and Mourato 2001). However, payment cards can introduce range bias arising from the lower and upper monetary level on the card. We use the pilot phase (see pilot report in Section 3.1) to test the range of payment options and address any potential range bias as best as we can.

2.4.3 Embedding effect and insensitivity to scope

Insensitivity to scope bias occurs where WTP is insensitive to the extent of the proposed change, such that stated welfare measures do not vary proportionally with the scope of the provided benefit. Similarly, inconsistency in WTP values has been observed where respondents are willing to pay the same amount for a set of goods (such as free access to all museums in the country) as for separate components (access values for individual museums), or are willing to pay the same amount for very different quantities of the same good (Bateman and Langford 1997; Hausman 2012; Kahneman and Knetsch 1992). Termed the embedding effect, or part-whole bias, this has considerable policy implications where evaluations are sensitive to the composition and quantity of goods

[15] Cheap talk script is a survey technique designed to reduce hypothetical bias in WTP estimates by reminding respondents of their budget constraints and availability of alternative goods, in order to make WTP values incentive compatible.

employed in the analysis. This is potentially a major concern in the present study, given that there are hundreds of museums in England and we are only valuing a maximum of three (and one at a time). Sequencing effects also matter as people are made aware of substitute sites they might value (discussed under ordering, see below).

2.4.4 Order Effects

Order effects are defined where responses to a given question vary in a theoretically unanticipated manner according to the positioning of that question relative to others in the survey instrument (Powe and Bateman 2003). We note that the valuations undertaken in this study are mutually exclusive. As such, the museum-level WTP values are not additive, and people are only asked to pay for one museum at a time. There would therefore be no budget constraint imposed for each subsequent museum. However, acting against this, a depressing effect introduced by question order could arise because people are made aware of substitute sites they also value. We test for order effects where respondents answer more than one museum visitor WTP question (sites are presented randomly, but their order is recorded for the purpose of sensitivity analysis). This gives an indication of any order effects that might exist within museum visitor use WTP values. The ability to test for order effects and control for them means that they should not in practice invalidate our survey design.

As the order in which visitor WTP is elicited is randomised within the survey the effect of survey order in the calculation of mean WTP should not be statistically significant across the whole sample of museum visitors.[16] However, in some cases a higher proportion of respondents have visited some museums compared with others.[17] This means that across the sample, certain museums may have a higher probability of being first in order (through higher sample size and more instances and visits to that museum only). We test in a sensitivity analysis whether this introduces significant differences in WTP through order effects.

Order effects may also arise in the museum non-visitor study groups. The WTP elicited for non-visited museums may be affected by the number of museum visit use WTP questions asked previously in the survey, compared to those who have not visited any of the four museums (and answer only the museum non-visitor non-use WTP question).

Potential order effect bias is introduced through the complex structure of the survey. These issues are not present in standard CV surveys where respondents are asked to value only one good or service. This is one of the trade-offs introduced through the specific design of the survey instrument and the cost constraints on sample size. In Section 3.5 we test whether these biases do in fact arise, and the success of bias reduction strategies like order randomisation.

[16] Malhotra (2008) shows how randomising the response order in questionnaires can reduce the impact of survey bias arising from survey satisficing where respondents are in a rush to complete the survey.

[17] For instance, the data shows that the National Railway Museum has been visited more than the other three museums. Randomisation of all museums visited within each survey means that the NRM is more likely to have come first in any valuation questions. We test whether this gives rise to an order effect in the valuation for the NRM.

2.5 Benefit transfer

There are two main approaches to BT in the literature (Brouwer 2000). The first is based on a transfer of a known benefit to another site; the second on the transfer of a function containing characteristics of the users and non-users of a site, as well as the characteristics of the site and possibly the study methodology, and how much these characteristics are associated with the valuation of the cultural and heritage goods in question. The first can be split into two sub-approaches, a straightforward value transfer and a transfer weighted by the relative incomes of the user and non-user groups.

We test these approaches to assess the validity of BT within our four case study institutions. The key element of the BT test is an analysis of the transfer error, i.e. the difference between the transferred value, and the value we estimate. To do this we use one of the sites in the study as a "policy site" and the others as the "study" site. In this section, we summarise these approaches. Section 4 sets out the approaches in more detail.

2.5.1 Transfer WTP on an institution by institution basis: unit value transfer and adjusted unit value transfer

This method is based on single point transfer estimates. Unit value transfer methods test the equality of mean WTP values at the policy site and the study site.

(i) ***simple unit value transfer***, where a single point estimate of benefit (e.g. mean WTP) is taken from one or more study sites and applied to the new policy site under the implicit assumption that the good and the socio-economic characteristics and preferences of the population are homogeneous between the study sites and the policy site:

Equation 5

$$\widehat{\boldsymbol{WTP}}_{\boldsymbol{p}} = \overline{\boldsymbol{WTP}}_{\boldsymbol{s}}$$

where $\widehat{WTP}_p$ is the predicted (average) WTP at the policy site and $\overline{WTP}_s$ is the average WTP at the study site(s); or the

(ii) ***Adjusted unit value transfer***, where the transfer accounts and controls for differences in conditions between the policy and study sites. This method usually focuses on differences in respondents' income, which could affect WTP estimates between two sites:

Equation 6

$$\widehat{WTP}_p = \left(\frac{\bar{Y}_p}{\bar{Y}_s}\right)^e \overline{WTP}_s$$

where $\bar{Y}_p, \bar{Y}_s$ is the average household income at policy and study sites, respectively, and e is the elasticity of the marginal utility of income with respect to WTP. We assume, as per the Green Book, that this equals 1 (i.e. e = 1).[18]

2.5.2 Value Function Transfer: Transfer adjusted WTP from pooled data

(iii) ***Benefit function*** (Desvouges et al. 1992a; Kaul et al. 2013; Loomis 1992b) representing the relationship between WTP and a number of explanatory variables. The researcher transfers the entire benefit function estimated at the study site(s) to the policy site, where it is adapted to fit the characteristics of the policy site (such as socio-economic characteristics and other measurable characteristics that systematically differ between the policy and study sites). The tailored benefit function is then used to predict the benefits for the policy site (Rosenberger and Loomis 2003):

Equation 7

$$\widehat{WTP}_{ip} = b_0 + b_1 Q_p + b_2 C_p + b_3 A_p + b_4 S_{ip}$$

where $\widehat{WTP}_{ip}$ is the predicted willingness to pay of individual i for policy site p; Q_p is the change in provision of the cultural good/service at site p; C_p is the characteristics of the good at site p; A_p is the availability of substitute sites for site p; and S_{ip} are the socio-economic characteristics of individual i at site p. The coefficients $b_0, \dots, b_4$ are obtained from the WTP function estimated at the study site (Equation 7 is estimated for the study sites whereby the subscripts p become subscripts s). Under this approach, more information about the site and population can be transferred and so the transfer errors are likely to be lower than the other two methods (Brouwer and Spaninks 1999). On the other hand, this approach is more data-intensive and requires availability of a range of demographic and possibly attitudinal/behavioural variables that are part of the WTP function, in each site.

Since for policy decisions, we are interested in an average WTP for a site, knowing the WTP per individual is not required. For this reason, we can average out Equation 7 across individuals:

Equation 8

$$\widehat{WTP}_p = b_0 + b_1 Q_p + b_2 C_p + b_3 A_p + b_4 \bar{S}_p$$

where now $\bar{S}_p$ is a set of the average socio-economic characteristics of individuals at site p; and the remaining notation is the same as in Equation 7. Equation 8 highlights the fact that individual-level data from the policy site are not necessary in order to predict average WTP. Rather, information on the average characteristics of the policy site is sufficient and this may be held by the policy site itself without the need for any further primary data collection.

In our study Equation 8 is obtained iteratively for each museum. Out of the four museums in the study we select a sub-set of three museums (which become the study sites) and

[18] Alternatively, the elasticity of the marginal utility of income could be estimated using data from the study site – this would be more in the spirit of the function transfer approach discussed below in the text.

estimate a benefit function on pooled data from these three study sites. The omitted fourth museum then becomes the policy site and characteristics from the omitted museum are plugged into Equation 8 to predict WTP at the policy site (Bateman et al. 2011). Each of the four museums in our study has "its turn" as a policy site and so the above process is conducted four times omitting a different museum each time which then becomes the policy site for that particular iteration of the study. We therefore predict WTP values for each of the four museums based on pooled benefit functions from the other three museums.

2.5.3 Transfer error testing

A number of transfer tests have been proposed to test the predictive power of BT. The statistical validity of benefit transfer is based on the assumption that value estimates are statistically identical across study and policy contexts. In other words, the values estimated for the pooled study sites should not be significantly different from the policy site. This difference, known as transfer error, is measured in two ways.

First, we calculate the percentage difference between the observed and the predicted WTP value. What is an acceptable transfer error and whether the transfer is still informative depends on the intended policy use of the transferred estimates, and the corresponding accuracy required (Brookshire and Neill 1992; Desvouges et al. 1992b). Here, we compare estimates of transfer error to established ranges within the literature (Mourato et al. 2014; Ståle Navrud and Ready 2007). Ready and Navrud (2006) reviewed intra and cross-country benefit transfer studies and found that the average transfer error was in the range of 20% to 40%, while individual transfers had errors as high as 100-200%, particularly when involving complex goods. For the purpose of testing we apply a threshold of maximum 40% transfer error to all individual transfer errors.

Second, we test the statistical difference between observed and predicted WTP in each case using student's t-tests. The acceptable threshold of statistically significant transfer error is not clearly set in the benefit transfer literature. For the purposes of transfer testing in this study we deem transfer errors to be acceptable if differences in observed policy site and pooled study sites WTP values are statistically insignificant in at least three of the four cases. Given the lack of guidance from the literature, we place more weight on transfer tests which produce errors below the 40% transfer error threshold proposed by Ready and Navrud (2006).

For use values across sites and populations, we test three hypotheses for the three BT methods outlined in Equations 5-7 (Table 2.2). For the transfer of non-use values across sites for the same general population we test only hypothesis 1 given the weaker explanatory power of the value functions (see Section 1.6.2).

Table 2.2 Benefit transfer tests employed

BT APPROACH	T-TEST HYPOTHESIS
UNIT TRANSFER	
Simple	$H1: \overline{WTP}_p = \overline{WTP}_s$ Null hypothesis: equivalence of observed mean policy site WTP and mean pooled study site WTP.
Adjusted	$H2: \frac{1}{a_p}\overline{WTP}_p = \frac{1}{a_s}\overline{WTP}_s$ where $a_i = (\bar{Y}_i)^e$ for $i = p, s$ Null hypothesis: equivalence of observed mean policy site WTP and mean pooled study site WTP, adjusted for income difference between policy and study site.
FUNCTION TRANSFER	
Function	$H3: \overline{WTP}_p = b \cdot \bar{X}_p$ Null hypothesis: equivalence of observed mean policy site WTP and mean predicted pooled study site WTP.

Notes: *$\overline{WTP}_p$, $\overline{WTP}_s$ = average WTP at policy (p) and study (s) sites; $\bar{Y}_p, \bar{Y}_s$= average household income at policy and study side respectively; $e = 1$; b = coefficients obtained from WTP function estimated at study sites; and $\bar{X}_p$ = average characteristics of the policy site. For simple and adjusted unit transfer approaches, we use the equivalent of a two-sample unpaired t-test with unequal variances for weighted data, for the function transfer approach we use a paired t-test.*

Hypothesis H1 tests the equality of mean WTP values at the **policy site** and the **study site.** Alternatively, average values from multiple study sites can be used, which is our approach here.

Hypothesis H2 tests the equality of *adjusted* mean WTP values at the policy site and the study site (or pool of study sites), adjusting for differences in any relevant characteristics. Accounting for differences in income is the most common adjustment, and is the approach we use here for use values.

Hypothesis H3 tests the transferability of a *pooled benefit function*, which is obtained after pooling the datasets from the study sites (excluding the policy case in each case) and estimating a WTP function for the pooled dataset. Specifically, H3 tests the equality of the observed mean WTP at the policy site and the predicted mean WTP for the policy site, using the estimated parameter coefficients of the pooled WTP function and the values of predictor variables observed at the policy site.

The accuracy of transfers (either unit or function transfers) is assessed by estimating the respective transfer errors, as follows:

Equation 9

$$TE = \left(\frac{\widehat{WTP_p} - \overline{WTP_p}}{\overline{WTP_p}} \right) \times 100$$

3 Contingent Valuation Results

3.1 Pilot survey

We implemented a pilot survey on 22nd February 2017 using a quota-based sample of 71 online panel respondents. Debrief questions were asked to identify potential problem areas in understanding or in survey design prior to going into the field for the final survey. The pilot survey also allowed us to test the range of WTP values provided in the payment cards, and amend payment card options if required. The pilot was performed under identical conditions to the final survey. The majority of respondents indicated that the survey length was okay or short (95%), that they did not find the survey difficult (96%), and that they had enough information on the purpose and aims of the survey (93%).[19] A full report on the pilot survey is available in Annex A5.

3.2 Implementation

We conducted an online survey on a total sample of 1,587 respondents (27th February – 30th March 2017), comprising both museum visitors and non-visitors.

We report on the results as they relate to the two key sub-samples (museum visitors and museum non-visitors) separately. Section 3.3 reports on the results for the visitors of the four museums, while Section 3.4 reports results for the museum non-visitors.

We test the survey internally to ascertain a realistic time required to give meaningful responses (those who complete the survey faster are termed "speedsters" in the CV literature: for further discussion, see Campbell et al. 2017). Due to the complex nature of the survey design, thresholds for exclusion of speedsters are set dependent on the number of WTP questions answered.[20] In the event, 65 surveys are excluded from the analysis on the basis of speedster responses.

[19] We subsequently made edits to the museum selection section of the survey with the inclusion of follow-up questions on the exact date that respondents visited the institutions selected (year and season), with additional options for 'Before 2014' and 'Never Visited Before'. The purpose of these follow-up questions was to reduce the probability of overstatement of visit frequency on the part of respondents, as identified in previous studies in the cultural sector (Bakhshi et al. 2015).

[20] Internal testing led to the following thresholds for speedsters based on number of WTP questions answered: One WTP question = 3:00; two WTP questions = 3:30; three WTP questions = 3:50; four WTP questions = 4:00. These thresholds take into account the fixed time resources required for common sections of the survey (Sections 1, 2, and 4, asked to all), and the time required for each additional WTP section.

3.3 Results: Museum visitors

3.3.1 Socio-demographics

Table 3.1 and Table 3.2 summarise the key socio-demographic characteristics across the four museum visitor study groups. For comparison, Table 3.1 reports socio-demographic results without weighting and Table 3.2 with weighting. Annex A6.1 & Annex A6.5 provide the weights used for the museums and a comparison with the population in general.

In all subsequent tables in this section (Table 3.3 onwards) we report only museum visitor-weighted figures for WTP estimates.

The unweighted socio-demographic characteristics for the museum visitor survey samples show that a higher proportion of respondents are female across all four museum visitor groups. The average age ranges from 39-49. We weight the sample of each museum to make it representative of all visitors for that museum. To that end, we source information from each of the museums on their total number of visitors and their breakdown by age and gender.

The unweighted socio-demographic characteristics show that average annual household income ranges from £40k to £44k. Between 44% (NRM) to 57% (Oxford Ashmolean) are university educated, while the majority across all four museums are in employment (57%-69%), married/with partner (54%-63%) and in good health (76%-85%). The highest proportion of museum visitors living in London exists within the Great North sample (29%), while the lowest exists within the NRM sample (15%). Between 30% (World Museum) and 43% (Great North Museum) were members of a cultural, conservation, environmental or other organisation.

Table 3.1 Museum visitor socio-demographic characteristics (unweighted)

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Male % (n/N)	37.1% (98/264)	34.0% (96/282)	45.1% (179/397)	27.8% (70/252)
Age (mean)	39.6	42.0	49.0	47.9
Household annual income (£, mean)	43,647	40,406	41,072	42,938
Dependent children under 16 years % (n/N)	56.1% (148/264)	45.4% (128/282)	31.5% (125/397)	28.6% (72/252)
Married/with partner % (n/N)	62.9% (166/264)	54.3% (153/282)	61.0% (242/397)	53.6% (135/252)
University education % (n/N)	52.3% (138/264)	53.9% (152/282)	43.6% (173/397)	57.1% (144/252)
In employment (full-time, part-time, self-employed) % (n/N)	68.9% (182/264)	66.7% (188/282)	57.4% (228/397)	65.9% (166/252)
Living in London % (n/N)	29.2% (77/264)	19.9% (56/282)	14.9% (59/397)	21.4% (54/252)
Health (good, very good, excellent) % (n/N)	79.2% (209/264)	81.9% (231/282)	75.8% (301/397)	84.5% (213/252)

Member of a cultural, conservation, environmental or other organisation % (n/N)	43.2% (114/264)	30.1% (85/282)	35.3% (140/397)	39.3% (99/252)
Visited more than one of the museums in past 3 years (%)	51.9% (137/264)	47.9% (135/282)	53.9% (214/397)	34.1% (86/252)

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Notes: Gross annual household income; averages computed using the midpoints of the income and age categories.

Table 3.2 Museum visitor socio-demographic characteristics (weights applied)

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Male (%)	54.7%	42.8%	53.0%	47.1%
Age (mean)	40.0	41.9	47.2	47.7
Household annual income (£, mean)	40,852	39,808	41,304	45,105
Dependent children under 16 years (%)	50.1%	42.1%	36.0%	29.2%
Married/with partner (%)	55.1%	51.7%	59.8%	51.9%
University education (%)	50.9%	49.4%	45.5%	58.9%
In employment (full-time, part-time, self-employed) (%)	67.2%	63.5%	61.5%	70.0%
Living in London (%)	22.8%	18.1%	18.6%	25.0%
Health (good, very good, excellent) (%)	80.8%	82.0%	76.3%	87.7%
Member (heritage, conservation, environmental org) (%)	39.9%	29.9%	33.9%	35.4%

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Notes: Gross annual household income; averages computed using the midpoints of the income and age categories. Weights are based on the breakdown by age and gender (see Annex A6).

3.3.2 Museum visits

Table 3.3 summarises information about visits to each of the four museums within the museum visitor study categories (defined as visitors in the past three years). The results show that between 72% and 89% of the museum visitors report having visited within the past two years and between 40% and 52% having visited in the last 12 months. On average, respondents have visited the museums they give use values for between 3 and 4 times in their lifetime. Although most visitors are from outside the city in which the museum is located, there is some variation across museums. In our sample, the Great North Museum has the highest proportion of respondents living in the same city as the museum (28.2%), followed by the Liverpool World Museum (18.4%), while the NRM has the lowest proportion (4.8%), which may suggest that the NRM serves a national audience, willing to travel from other parts of the country, while the other museums attract more local audiences. We note, however, that Liverpool and Tyne and Wear are more heavily populated areas, while York and Oxford may have wider regional

hinterlands which are not captured in the city location question. We explore the differences in WTP values by geographical region in Section 3.3.5.

Only a minority of respondents express high familiarity (very or extremely familiar) with the information presented in the survey on each of the museums, ranging from 39.2% (NRM) to 25.5% in the case of the World Museum.

Table 3.3 Museum visitor usage information (visitor weights applied)

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Total number of visits in lifetime (mean / median)	3.92 / 2	3.46 / 2	3.44 / 2	3.36 / 2
Visited the institution in the last 2 years (%)	81.1%	88.3%	77.5%	72.0%
Visited the institution in the last 12 months (%)	40.8%	51.5%	41.7%	44.7%
Resident in city (%)	28.2%	18.4%	4.8%	8.0%
Familiarity with information (very or extremely familiar) (%)	36.7%	25.5%	39.2%	28.1%

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Sample weighted by visitor weights.

3.3.3 Attitudes

Attitudes towards culture and museums are depicted for each museum visitor sub-sample in Table 3.4. The table shows a high rate of participation and engagement with culture and heritage. Across all four museums, between 90% and 96% of respondents have visited other museums and galleries in the last 12 months. Between 80% and 87% had been taken to museums and galleries as a child.

Between 34% and 37% of respondents place arts, culture & heritage amongst the 5 top priorities for public spending. These percentages are higher than the 26-27% of museum visitors surveyed in the Natural History Museum and Tate Liverpool surveys in our earlier study (Bakhshi et al. 2015). This result accords with our expectations as the question used in the latter study was more stringent, asking about the top 3 priorities for Government spending (Bakhshi et al. 2015, p. 33; 74).

In terms of agreement questions, we find that a large majority of respondents across all museum visitor groups agree or strongly agree that preserving museums for the appreciation of current and future generations is important (78%-85%).

Table 3.4 Museum visitor attitudes towards culture and museums (weights applied)

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Visited other museums and galleries in last 12 months (%)	92.0%	93.5%	90.7%	95.2%

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Taken to museums and galleries as a child (%)	83.1%	86.9%	80.4%	84.3%
Arts, culture & heritage amongst the 5 top priorities for public spending (%)	34.6%	36.6%	36.6%	36.2%
Preserving museums for the appreciation of current and future generations is important to me (%)	78.1%	78.8%	84.5%	84.6%
Visiting museums increases one's wellbeing (happiness) (%)	65.5%	74.8%	79.9%	83.6%
Museums have a value even for those who do not visit (%)	66.7%	65.6%	73.4%	65.3%
There are more important things to spend my money on than museums (%)	35.9%	33.0%	31.2%	31.8%

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Sample weighted by visitor weights.

3.3.4 WTP summary statistics (use values)

Table 3.5 shows the proportion of museum visitors who indicate that they are in principle willing to pay an entry fee to access the museum under study.

We see that in all cases, a high proportion of museum visitors are in principle willing to pay (Yes or Maybe) an entry fee to access the museum. 95.5% are willing to pay in principle for the Great North Museum, 93.6% for Liverpool World Museum, 94.5% for NRM and 91.3% for Ashmolean museum. Table 3.5 shows that the share of visitors that are not willing to pay in principle is conversely low, between 4.5% for the Great North Museum, and 8.7% for the Ashmolean. These shares are lower than found in other studies: e.g. 15%-20% in the Natural History Museum and Tate Liverpool case studies (Bakhshi et al. 2015).

Table 3.5 Museum visitor willingness to pay in principle

WTP in principle	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Yes	64.4% (170/264)	63.8% (180/282)	71.3% (283/397)	67.5% (170/252)
Maybe	31.1% (82/264)	29.8% (84/282)	23.2% (92/397)	23.8% (60/252)
No	4.5% (12/264)	6.4% (18/282)	5.5% (22/397)	8.7% (22/252)

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Sample weighted by visitor weights.

Table 3.6 shows the mean and median WTP of museum visitors, via an entry fee, to access the relevant museums. The mean WTP of visitors is between £7.79 (median £5.50) for the Great North Museum and £6.01 (median £4.50) for the World Museum. The mean WTP for the National Railway Museum is £6.86 (median £5.50) and £7.08 (median £5.50) for the Ashmolean Museum. These WTP values are comparable to use values estimated in previous studies for the Natural History Museum (£6.65) and Tate Liverpool (£10.83) (Bakhshi et al. 2015).

The proportion of valid zero WTP answers is low (both from those that stated they were not in principle willing to pay, and those that said they might be willing to pay in principle but then decided not to), between 5% and 11%. This is suggestive that the scenario presented was valued and realistic, and the range of payment amounts offered was credible and affordable.

Table 3.6 Museum visitor mean and median use WTP (entry fee, weighted). Mean and median values calculated with inclusion of 'No' at payment principle (coded £0).

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
No. of visitors	264	282	397	252
Mean (standard error)	£7.79 (£0.57)	£6.01 (£0.39)	£6.86 (£0.41)	£7.08 (£0.86)
95% CI low	£6.66	£5.23	£6.05	£5.39
95% CI high	£8.91	£6.78	£7.66	£8.77
Median	£5.50	£4.50	£5.50	£5.50
Max	£75.00	£67.50	£57.50	£75.00
Zeros (including those not WTP in principle)	5.7% (15/264)	7.1% (20/282)	6.0% (24/397)	10.3% (26/252)
Zeros, of those WTP in principle	1.1% (3/264)	0.7% (2/282)	0.5% (2/397)	1.6% (4/252)

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). All WTP values are calculated as the midpoint interval between the selected payment amount in the payment card and the next highest response on the payment card (except £0 bids). Summary statistics calculated with inclusion of 'No' at payment principle (coded £0). Sample weighted by visitor weights.

3.3.5 WTP by socio-demographic group

When we disaggregate museum visitors' use WTP into socio-demographic groups, we see that those with a higher income are willing to pay more to visit a museum (Annex A6.3). Those from the North-East do not have significantly higher mean WTP for the Great North

Museum (£4.33) when compared with those from other regions (Greater London provides the highest mean WTP, at £13.46). Those from the North-West do not have significantly higher mean WTP for the World Museum (£4.20) than those from other regions (Greater London again provides the highest mean WTP, at £10.41). Those from Yorkshire and the Humber do not have significantly higher mean WTP for the NRM (£5.06) than those from other regions (in this case, the West Midlands provides the highest mean WTP, at £11.26). Those from the East or South-East of England do not provide a higher mean WTP for the Ashmolean (£8.41 and £5.23 respectively) compared with those from other regions (Greater London provides the highest mean WTP, at £10.42). The higher mean WTP for visitors from Greater London for three of the four museums is suggestive of an income effect operating as a driver of WTP. We note, however, the small number of observations for some of the sub-groups, so these figures should only be seen as indicative.

3.3.6 Validity testing: WTP determinants

As described in Section 2.3.2. we check the theoretical validity of our results by testing if museum visitors' WTP (as an entry fee to access the relevant museums) is associated with different socio-demographic, and behavioural (number of visits i.e. usage and knowledge) and attitudinal factors that are theoretically expected to drive WTP.

Our choice of independent variables follows the recommendations of Bateman et al. (2002), which is common practice in modern applications of CV. In particular, we include a range of standard socio-demographic variables (i.e. gender, age, children, ethnicity, education and income), variables relating to previous usage of visited museums, and relevant attitudinal variables (e.g. membership in heritage associations; attitudes towards museums and heritage and public spending on culture). Annex Table A6.4 summarises the variables used.

Table 3.7 shows the results for museum visitor use WTP in terms of their willingness to pay an entry fee to access museums, controlling for a range of factors.

Table 3.7 Factors associated with museum visitors WTP, as an entry fee to access each museum

	Great North Museum	World Museum	National Railway Museum	Ashmolean	Pooled
Female	3.579***	0.684	-0.809	-1.942	0.090
Log age, using age midpoint	-0.451	-0.759	-0.064	-2.401	-0.408
Log income, using income midpoints	1.174	1.696***	0.998**	1.524	1.425***
BAME	-2.240	1.091	-0.114	3.110	0.329
Degree and above	-1.929*	0.359	-0.242	1.640	-0.040
Married, civil partner, cohabiting	-1.786	0.657	-1.413	2.024	-0.433
Employed (FT, PT, SE)	1.738	0.713	1.416**	-1.425	0.417
Dependent children	1.527	0.424	-0.080	2.113	0.720
Resident of city of museum	5.297***	1.379	7.506*	-1.448	2.317**
Log geodesic distance to museum (miles)	2.074***	0.963***	0.884*	1.509	1.214***
Num visits museum (lifetime)	0.566***	0.120	-0.143	0.720**	0.218**

Last visit to museum=2014	0.000	0.000	0.000	0.000	0.000
Last visit to museum=2015	-0.355	-0.725	1.220	1.056	0.208
Last visit to museum=2016	-0.648	-0.575	0.138	-0.365	-0.209
Last visit to museum=2017	-1.541	-1.646	-0.431	2.490	-0.347
Visited cultural institution in past year (e.g. museum, gallery, historic building	-2.036	-2.260	-1.897	1.768	-1.753
Member of a cultural, conservation, environmental or other organisation	1.649	1.812**	2.069	-2.947	1.114
Familiar with information on museum: Very/Extremely	1.583	2.799***	2.493***	2.395*	2.126***
Arts, culture & heritage are a fiscal priority	0.974	0.607	1.110	-0.800	0.339
Preserving museums for current and future generations is important (agree/strongly agree)	1.024	1.855**	1.960**	3.408*	2.157***
Constant	-16.581*	-15.914**	-9.214	-14.937	-14.731***
Observations	221	254	363	237	1075
Adjusted R^2	0.254	0.213	0.185	0.197	0.158

*Notes: *** significance at <1%; ** significance at <5%; * significance at <10%. Reference group: for gender ref = male; for BAME ref = white; for education Degree and above ref = all qualifications under Degree; for Married, civil partner, cohabiting ref = other marital status; Employed (FT, PT, SE) ref = other employment status; for Dependent children ref = no children; for good-excellent health ref = poor/fair health; for London resident ref = rest of England; for Familiar with information on museum: Very/Extremely ref = not at all – moderately familiar. Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Sample weighted by visitor weights. Notes: Gross annual household income; averages computed using the midpoints of the income and age categories. Heteroskedasticity-robust standard errors. Controls for entreaty script (not shown). Museum specific dummies insignificant in pooled regression. All VIF scores <2 in pooled regression.*

We assess the models in Table 3.7 according to whether WTP is associated with theoretically consistent drivers of value. In best-fit pooled regression (Table 3.7, last column) and in two of the museum visitor models, income is significantly and positively associated with mean WTP an entry fee for access (the World Museum and the National Railway Museum). This is consistent with previous CV studies. Employment is also significantly and positively associated with mean WTP in the case of the NRM.

We find that residents of Newcastle (where the Great North Museum is located) and York (where the National Railway Museum is located) have a significantly higher WTP than those who live in the rest of England. Economic values of culture are commonly found to exhibit distance decay, such that local amenities are more highly valued than similar amenities further away. This finding is also consistent with the higher values found for Tate Liverpool among residents in the North-West (Bakhshi et al. 2015).

Distance from the respondent's postcode to the museum is significantly and positively associated with WTP. In other words, the further a respondent had to travel to visit the museum, the higher their associated WTP. This is after controlling for local effects of museum residents. In other words, those who live in the city (at a short distance to where the museum is based) have higher values than those who live elsewhere in the country,

but within the non-resident population, WTP values increase with further distance travelled to it.

We include a number of indicators of cultural engagement within the validity testing model, to test our results against the theoretical assumption that those who are more engaged in culture (i.e. visit cultural institutions in their spare time or prioritise public spending on arts and culture) would value the museums they visit more highly (Noonan 2003). For the World Museum we find a significant and positive association between membership of cultural, conservation, environmental or other organisations, compared with those who are not members of any organisations. There is no significant association between prioritisation attached to public spending on arts, culture and heritage and mean WTP. However, in three of the museums (the World Museum, NRM, Ashmolean) we find a positive association with mean WTP for those who agree or strongly agree on the importance of preserving museums for current and future generations.

We also test for validity using information on respondents' previous usage of the museum. We would anticipate that those who have used a museum more will have a higher use value for the institutions. In the case of two museums (Great North Museum and Ashmolean) and in the pooled model, this is indeed the case. The year in which the last visit took place is not, however, significantly associated with mean WTP in any of the four museums.

Familiarity with the information presented about the museum is also significantly and positively associated with mean WTP in the pooled model and in three of the museum sites (World Museum, National Rail Museum and Ashmolean).

In addition, indicators of cultural engagement are significant within the pooled regression: respondents who are extremely/very familiar with information on the museum have a significantly higher mean WTP than those who are not so familiar. We also find that respondents who agree or strongly agree that preserving museums for current and future generations is important have a significantly higher mean WTP in the pooled regression.

Sensitivity testing of follow-up certainty questions is reported in Annex A3. Tests of model fit (Adjusted R2) range between 0.19 and 0.25 for the individual museum models, and 0.16 for the pooled regression (perfect model fit would be measured as 1.0). This compares favourably with validity tests run using similar CV instruments to elicit WTP values in the cultural sector (in the Natural History Museum (R2=0.11-0.19) and Tate Liverpool (R2=0.06-0.15); see Bakhshi et al. 2015). These measures of model fit are therefore acceptable, considering that WTP is likely to be influenced by unobserved factors for which we do not have data, and which cannot therefore be included in the models.

3.3.7 Summary: Museum visitor use-value analysis

The mean WTP of museum visitors is between £7.79 and £6.01 across the four study museums. This is comparable to use values estimated in previous studies (Bakhshi et al. 2015).

The results accord with theoretical expectations, being positively and significantly associated with income, past usage of the museum under study, and indicators of cultural engagement (within the pooled museum visitor regression model and across at least three museum study sites).

The results indicate that the WTP estimates appear to be valid and therefore fulfil one of the conditions to allow reliable benefit transfer (Johnston et al. 2015, p. 125).

3.4 Results: Museum non-visitors

3.4.1 Socio-demographics

Table 3.8 and Table 3.9 summarise the key socio-demographic characteristics across the four museum non-visitor study groups. For comparison, Table 3.8 reports socio-demographic results without weighting and Table 3.9 with weighting. In all subsequent tables in this section we report nationally representative weighted figures (gender and age, APS, 2015-2016).

The unweighted socio-demographic characteristics for the museum non-visitor survey samples show that a higher proportion of respondents are female across all four museum visitor groups. The average age ranges from 47-50. When we apply weighting, the proportion of males is corrected upwards to 49% (Table 3.9). Average age is also adjusted slightly to 47 years old.

The unweighted socio-demographic characteristics show that average annual household income ranges from £34k to £38k. Between 35% (Ashmolean) to 43% (World Museum) are university educated, while over half of the non-visitors for all four museums are in employment (50%-57%). The highest proportion of museum non-visitors living in London is found within the World Museum sample (20%), while the lowest is found within the Ashmolean sample (18%). Between 23.1% (Great North Museum) and 28.4% (World Museum) are members of a cultural, conservation, environmental or other organisation.

Annex A6.5 provides the general population weights used for museum non-visitors.

Table 3.8 Museum non-visitor socio-demographic characteristics (unweighted)

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Male (% n/N)	38.7% (151/390)	43.8% (168/384)	34.9% (123/352)	43.1% (180/418)
Age (mean)	47.9	49.5	47.2	46.6
Household annual income (£, mean)	35,763	38,120	34,011	34,346
Dependent children under 16 years (% (n/N))	27.9% (109/390)	25.8% (99/384)	27.8% (98/352)	31.8% (133/418)
Married/with partner (% n/N)	48.2% (188/390)	51.6% (198/384)	45.2% (159/352)	48.8% (204/418)

University education (% n/N)	36.7% (143/390)	43.2% (166/384)	40.9% (144/352)	35.4% (148/418)
In employment (full-time, part-time, self-employed) (% n/N)	50.5% (197/390)	55.2% (212/384)	54.5% (192/352)	57.2% (239/418)
Living in London (% n/N)	18.2% (71/390)	19.8% (76/384)	19.6% (69/352)	17.9% (75/418)
Member of a cultural, conservation, environmental or other organisation (% n/N)	23.1% (90/390)	28.4% (109/384)	24.4% (86/352)	25.1% (105/418)

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Sample weighted by gen. pop. Weights (age and gender). Notes: Gross annual household income; averages computed using the midpoints of the income and age categories.

Table 3.9 Museum non-visitor socio-demographic characteristics (weighted)

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Male (%) (weighting variable)	48.9%	48.9%	48.9%	48.9%
Age (mean) (weighting variable)	47.2	47.2	47.2	47.2
Household annual income (£, mean)	35,159	36,726	32,482	33,664
Dependent children under 16 years (%)	28.4%	27.0%	25.2%	32.0%
Married/with partner (%)	45.5%	47.4%	41.1%	46.0%
University education (%)	36.7%	43.5%	39.7%	35.3%
In employment (full-time, part-time, self-employed) (%)	50.7%	56.0%	52.6%	55.0%
Living in London (%)	18.8%	20.1%	19.0%	21.3%
Member (heritage, conservation, environmental org) (%)	20.6%	27.7%	23.8%	25.0%

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Notes: Gross annual household income; averages computed using the midpoints of the income and age categories.

3.4.2 Attitudes

Attitudes towards culture and museums are depicted for each sub-sample of museum non-visitors in Table 3.10. The table shows a high rate of participation and engagement with culture and heritage. Across all four museums, between 72% and 81% of respondents had visited other museums and galleries in the last 12 months. Between 77% and 84% had been taken to museums and galleries as a child.

Between 27% and 33% of respondents place arts, culture & heritage amongst their top 5 top priorities for public spending. As with the visitor results, these percentages are slightly higher than the 26-27% of museum visitors surveyed in the Natural History Museum and Tate Liverpool surveys (Bakhshi et al. 2015). We find that the highest

proportion of respondents across all museum non-visitor groups agree or strongly agree that preserving museums for the appreciation of current and future generations is important (28%-40%).

Table 3.10 Museum non-visitor attitudes towards culture and museums (weights applied)

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Visited other museums and galleries in last 12 months (%)	75.4%	79.9%	68.7%	75.5%
Taken to museums and galleries as a child (%)	80.8%	79.2%	83.8%	79.5%
Arts, culture & heritage amongst the 5 top priorities for public spending (%)	27.7%	27.2%	33.1%	29.0%
Preserving museums for the appreciation of current and future generations is important to me (%)	34.2%	39.3%	28.7%	32.4%
Visiting museums increases one's wellbeing (happiness) (%)	24.9%	30.1%	20.5%	20.4%
Museums have a value even for those who do not visit (%)	25.3%	24.7%	19.2%	21.9%
There are more important things to spend my money on than museums (%)	8.2%	9.9%	6.2%	12.2%
Resident in the city (%)	1.6%	5.5%	3.3%	5.5%
Familiarity with information (very or extremely familiar) (%)	3.2%	8.9%	7.7%	7.4%

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Sample weighted by gen. pop. weights (age and gender).

3.4.3 WTP summary statistics (non-use values)

Table 3.11 shows the proportion of museum non-visitors who indicate that they are in principle willing to pay an annual donation to ensure that collections are adequately conserved and maintained, and presented in the best possible way for each institution.

We see that, in all cases, around 50%-55% of museum non-visitors are in principle willing to pay (Yes or Maybe) an annual donation. As such, the share of non-visitors who are not willing pay in principle is much higher (between 45%-50%) than for museum visitors (between 5%-9% for museum visitors). We would expect that those who have not previously used a cultural institution would be less likely to be willing to pay to support its work.

This share of non-users who would not be willing to pay in principle is high compared with the proportions found in the previous Natural History Museum study (22%) (a museum of international reputation), but in line with proportions not willing to pay in principle in the Tate Liverpool study (53%) (Bakhshi et al. 2015).

Table 3.11 Museum non-visitor willingness to pay in principle (weights applied)

WTP in principle	Great North Museum (%, N)	World Museum (%, N)	National Railway Museum (%, N)	Ashmolean Museum (%, N)
Yes	17.4% (68/390)	18.0% (69/384)	17.6% (62/352)	18.9% (79/418)
Maybe	37.2% (145/390)	35.2% (135/384)	32.7% (115/352)	33.5% (140/418)
No	45.4% (177/390)	46.9% (180/384)	49.7% (175/352)	47.6% (199/418)

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Sample weighted by gen. pop. weights (age and gender). Note: All No responses coded as £0 for WTP analysis.

Table 3.12 shows the mean and median WTP of non-users (non-visitors), via an annual donation, to support the maintenance and conservation of the museum's collection.

Among museum non-visitors, mean WTP ranges from £2.79 (median £1.30) in the case of the Great North Museum to £4.06 in the case of the Ashmolean Museum (median £0.40). Mean non-visitors WTP to the World Museum is £3.70 (median £1.30) and £3.30 (median 0.00) for the NRM.

As expected, the range of non-use WTP values among museum non-visitors is lower than the use WTP among museum visitors. As outlined in Section 3.3.6, we would expect that those who have used a cultural institution previously would have higher WTP to preserve it (Noonan 2003). Distributions of museum non-visitor WTP values are presented in Annex A2.

Table 3.12 Museum non-visitor mean and median non-use WTP (annual donation, weighted). Mean and median values calculated with inclusion of 'No' at payment principle (coded £0)

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Sample size	390	384	352	418
Mean (standard error)	£2.79 (£0.23)	£3.70 (£0.30)	£3.30 (£0.43)	£4.06 (£0.43)
95% CI low	£2.35	£3.10	£2.46	£3.21

95% CI high	£3.24	£4.29	£4.15	£4.91
Median	£1.30	£1.30	£0.00	£0.40
Max	£21.00	£52.50	£37.50	£75.00
Zeros (including those not WTP in principle)	48.2% (188/390)	49.0% (188/384)	54.3% (191/352)	49.3% (206/418)
Zeros, of those WTP in principle	2.8% (11/390)	2.1% (8/384)	4.5% (16/352)	1.7% (7/418)

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Sample weighted by gen. pop. weights (age and gender). Note: All WTP values are calculated as the midpoint interval between the selected payment amount in the payment card and the next highest response on the payment card (except for £0). Summary statistics calculated with inclusion of 'No' at payment principle (coded £0).

3.4.4 WTP by socio-demographic group (non-users)

When we disaggregate museum non-visitor WTP into socio-demographic groups, we see that those with a higher income are willing to pay more to visit a museum (Annex A6.6). Those from Yorkshire and the Humber have significantly higher mean non-use WTP for the NRM (£7.17) than those from other regions. However, those from the North-East do not have significantly higher mean non-use WTP for the Great North Museum (£3.17) compared with those from other regions (the East of England provides the highest mean non-use WTP, at £3.75). Those from the North-West do not have significantly higher mean non-use WTP for the World Museum (£3.94) than those from other regions (Greater London provides the highest mean WTP, at £5.11). Those from the East or South East of England do not provide a higher mean non-use WTP for the Ashmolean (£5.73 and £2.58 respectively) compared with those from other regions (Greater London provides the highest mean non-use WTP, at £7.27). The higher mean non-use WTP provided by Greater London for two of the four museums is again suggestive of an income effect operating as a driver, in this case of non-use WTP. As before, we note the small number of observations for some of the sub-groups so these figures should be seen only as indicative.

3.4.5 Validity testing: WTP determinants (non-users)

We investigate the theoretical validity of our results by establishing if non-visitor WTP (as a donation for the conservation of the collections of the museums) is associated with different socio-demographic, knowledge, and attitudinal factors that are theoretically expected to drive WTP (Bateman et al. 2002) (Annex Table A6.7).

Table 3.13 shows the results for factors associated with museum non-visitor WTP, as an annual donation, to ensure that collections are adequately conserved and maintained, and presented in the best possible way.

Table 3.13 Factors associated with non-visitor willingness to pay, as an annual donation for conservation of the collections

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum	Pooled
Female	0.092	0.722	-0.550	1.103	0.396
Log age, using age midpoint	0.189	-0.550	-1.296	-1.337	-0.948**
Log income, using income midpoints	-0.038	0.112	0.581	1.239***	0.582**
BAME	0.942	1.220	0.245	-0.455	0.883
Degree and above	0.000	0.109	-1.703**	0.096	-0.147
Married, civil partner, cohabiting	0.205	-0.637	-0.051	0.591	0.018
Employed (FT, PT, SE)	0.128	0.613	0.148	0.212	0.383
Dependent children	1.138**	-0.031	0.012	0.511	0.413
Resident of city of museum	-1.869	-0.927	3.014	0.418	1.601
Log geodesic distance	0.371	-0.719	0.260	-1.279*	-0.580*
Visited other cultural institution in past year (e.g. museum, gallery, historic building)	0.532	0.761	-1.479*	-0.778	-0.420
Member of a cultural, conservation, environmental or other organization	0.536	0.847	1.959**	-0.734	0.388
Familiar with information on museum: Very/Extremely	-0.109	5.781**	2.154	14.569***	6.375***
Arts, culture & heritage are a fiscal priority	0.843	0.674	3.170***	0.507	1.483***
Preserving museums for the appreciation of current and future generations is important (Agree/Strongly Agree)	1.011**	2.007***	0.700	2.161***	1.342***
Constant	-1.215	4.879	1.768	1.478	2.730
Observations	354	342	323	369	1388
Adjusted R2	0.124	0.137	0.139	0.372	0.193

*Notes: *** significance at <1%; ** significance at <5%; * significance at <10%. Reference group: for gender ref = male; for BAME ref = white; for education Degree and above ref = all qualifications under Degree; for Married, civil partner, cohabiting ref = other marital status; Employed (FT, PT, SE) ref = other employment status; for Dependent children ref = no children; for London resident ref = rest of England; for Familiar with information on museum: Very/Extremely ref = not at all – moderately familiar. Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Sample weighted by gen. pop. weights (age and gender). Notes: Gross annual household income; averages computed using the midpoints of the income and age categories. Heteroskedasticity-robust standard errors. Controls for entreaty script (not shown). Museum specific dummies insignificant in pooled regression. All VIF scores <2 in pooled regression.*

In three of the non-visitor models, income is not significantly associated with mean WTP. This contrasts with many previous CV studies of cultural institutions which find that individuals earning higher income are more likely to pay more to support the work of cultural institutions (Noonan 2003). However, it is not an uncommon finding when dealing with very small WTP amounts (e.g. our median WTP values vary between 0 and £1.3) that are unlikely to be constrained by income. When the four non-visited museums are pooled – providing greater sample size for analysis – we find that income is significantly and positively associated with non-use WTP in a way that is consistent with theoretical expectations.

Distance between museum non-visitor's homes and the museums is found to be significantly and negatively associated with mean WTP for the Ashmolean museum, and in the pooled model. In other words, mean non-use WTP is *lower* for those who live further away from the institution. These results indicate the presence of benefit distance-decay, which is common in WTP studies. It reflects the likely presence of closer substitutes for the types of non-use benefits under investigation. Nonetheless, living in the city where the museum is based is not significantly associated with WTP in any of the non-use models.

Non-use WTP is significantly associated with indicators of cultural engagement, which we would expect based on previous studies of cultural institutions. In one museum (the NRM) those who are members of cultural, conservation, environmental or other organisations have a significantly higher mean non-use WTP.

Those who rank spending on arts, culture & heritage among their top 5 fiscal priorities have a significantly higher mean non-use WTP in the case of one museum (National Railway Museum), while familiarity with information about the museum is significantly and positively associated with mean non-use WTP in the World Museum and Ashmolean.

Analysing the coefficients obtained from the best-fit pooled regression (Table 3.13, final column), we find that the age (log) of non-visitors is significantly and negatively associated with their WTP an annual donation, although only at the 10% level.

Sensitivity testing of follow up certainty questions is reported in Annex A3. Within the best fit pooled regression, we find that non-visitors who are extremely or very familiar with information on the museums have a significantly higher WTP an annual donation. Furthermore, there is a significant and positive association between non-visitors who rank spending on arts, culture & heritage among their top 5 fiscal priorities and their WTP to pay an annual donation in the pooled regression.

Tests of model fit (Adjusted R2) range between 0.12 and 0.37 for the individual non-visited museum models, and 0.19 for the pooled regression. Again, these results compare favourably with validity tests run using similar CV exercises (for instance, the Natural History Museum (R2=0.11-0.19) and Tate Liverpool (R2=0.06-0.15); see Bakhshi et al. 2015). These measures of model fit are therefore acceptable, considering that willingness to pay is likely influenced by unobserved factors which cannot be included in our model.

3.4.6 Summary: Museum non-visitor non-use value analysis

The mean non-use WTP for our four museums ranges from £2.79 to £4.06, which appears reasonable for annual donations.

Mean and median WTP is lower among non-visitors than for visitors in all cases. This is as we would anticipate, based on theoretical expectations that users should hold higher values for a good or service than non-users (Bateman et al. 2002), and given the difference in payment vehicles used (entry fee vs donation) and the scenarios being valued (access vs. conservation of collections).

Validity tests indicate that the signs of the coefficients associated with non-use WTP results broadly conform to expectations, with income and cultural engagement

significantly and positively associated with mean non-use WTP at least within the pooled non-use model.

The results indicate that in many cases the museum non-visitor non-use WTP results are associated with factors that are theoretically expected to drive WTP, such as income and cultural engagement. However, the factors which are associated with WTP are not consistent across museums. This suggests that we need more data to better investigate the relationships between non-user WTP for museums.

3.5 Strategies for bias minimisation

As outlined in Section 2.4.3, our survey design potentially introduces different biases due to its nested nature. Respondents may be asked to independently value different cultural institutions in sequence within the same survey. Respondents are identified according to the museums they report having visited and may answer different numbers of WTP questions depending on the institutions they have visited.

We outline below a number of potential bias associated with these features, and the set of *ex ante* and *ex post* corrections we introduce to deal with them.

Hypothetical bias caused by the hypothetical nature of the CV survey have been found to lead to an overstatement of what respondents would pay in reality (Cummings and Taylor 1999b; Landry and List 2007; Mahieu et al. 2012). A range of counteractive treatments or corrective adjustments can be made to address hypothetical bias. Counteractive (i.e., *ex ante*) treatments, are often employed through so-called entreaties in the survey text. Famously, Cummings and Taylor (1999) developed a *cheap talk entreaty* for reducing hypothetical bias, whereby a script describes the bias problem and a plea is made to respondents that they do not overstate their true willingness to pay. The evidence finds that cheap talk scripts reduce hypothetical bias (e.g., Aadland and Caplan 2006; Carlsson et al. 2005; Carlsson and Martinsson 2006; List and Lucking-Reiley 2000; Lusk 2003). Another entreaty is the *oath script,* which typically asks respondents to agree to promise that they will respond to questions or state values honestly. Our earlier study for the DCMS and AHRC (Bakhshi et al., 2015) effectively used both types of entreaties and we employ them here. In particular, we randomise the application of entreaty script over four arms of the study within the sample (no entreaty/cheap talk only/oath script only/cheap talk and oath script) to test their effectiveness in this policy setting.

Order effects, also known as 'sequencing effects', refer to the differences in respondents' WTP that result from the order (e.g. first, second or third) in which the museums are presented to them. Our tests of order effects within those who answered more than one museum visitor use WTP question indicate that the order in which museums are presented is not significantly associated with mean WTP (Annex A7). This improves our confidence that the randomised design of the survey (eliciting values for multiple but randomly presented museums from the same respondents within the same survey instrument) is a successful way to mitigate against potential bias arising from order effects. On this basis, we do not recommend that any further *ex post* corrections are needed to correct for order effects. We note, however, that the available sample sizes for testing the effect of question order among each of the museum samples is small. We would encourage further research into the effects of WTP question ordering on larger

samples in future CV studies that are designed to value multiple goods within the same survey instrument.

With respect to non-use WTP, it is not possible to test for the effects of WTP question ordering within this study due to the overriding influence of self-selection by those who are more culturally engaged into the sub-group of respondents who answer one or more previous WTP questions. Given the distinct cultural goods being valued, the different hypothetical scenarios, payment vehicles, and respondent samples used to estimate non-use WTP, however, we might speculate that any WTP question ordering biases on the estimated non-use WTP values would be small. But, in order to rule out such effects with certainty it would be necessary to develop a separate survey that elicited only non-use values in isolation.[21]

Recollection bias, refers to a systematic error caused by differences in the accuracy of the recollections of participants, regarding their experience at the museums. We minimise this bias through follow-up questions that ask respondents to verify exactly when they visited, excluding those who fall outside of the 3-year period (recall Section 2.1.1).

Over-representation of respondents answering multiple WTP questions for multiple institutions within the pooled sample. As outlined in Section 2.2.1, we weight our data to reflect the profile of visitors at the national level and limit the effects of over-representation of individuals answering multiple WTP questions by adjusting their weights to the number of use values provided. We note that the maximum number of WTP values that can be provided by one individual in the pooled museum visitor data is 3, and these responses make up only a small proportion of the 1,195 observations. Note that all respondents answer only one non-use question in the museum non-visitor pooled data.

3.6 Discussion of use and non-use-values

Table 3.14 summarises the main results from our study of the four museums: the Great North Museum, the World Museum, the National Railway Museum and the Ashmolean Museum. As with all economic valuation studies, care must be taken when using and interpreting the range of values estimated.

Visitor use value is measured as a WTP a per person entrance fee for a museum which is currently free to enter. Museum visitor WTP is weighted by museum-level weights based on statistics from each of the study museums, to ensure representativeness to the populations of relevance. We interpret this value as a use value, but there could be some non-use elements within this value. We note the conceptual and practical difficulties of separating direct use and non-use value among visitors (Bakhshi et al. 2015). We cannot,

[21] An alternative might be to pool use and non-use WTP questions and randomise their order within the survey. However, previous studies on nested CV survey designs have shown that order effects are heightened when an inferior good is presented prior to a superior good (Powe and Bateman 2003). This effect may operate in a similar way for use WTP (theoretically expected to elicit higher values) and non-use WTP (theoretically expected to elicit lower values).

therefore, exclude the possibility that visitor use values capture some element of non-use value, and vice versa.

The non-use values are captured in the same survey through an annual donation to ensure that collections are adequately conserved and maintained, and presented in the best possible way for each institution. We apply nationally representative weights for England to ensure representativeness to the general population.

The values elicited in Table 3.14 therefore refer to different scenarios, payment settings, and population groups. When used for policy purposes it is important to identify clearly which value is most appropriate, depending on whether it refers to visitors or non-visitors, as well as the influence that the payment mechanism (entry fee or donation) has on the size of the value elicited.

Table 3.14 Summary of use and non-use Willingness to Pay values

Visitor WTP or Use value	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Mean (standard error)	£7.79 (£0.57)	£6.01 (£0.39)	£6.86 (£0.41)	£7.08 (£0.86)
Median	£5.50	£4.50	£5.50	£5.50
Non-visitor WTP or Non-use value	**Great North Museum**	**World Museum**	**National Railway Museum**	**Ashmolean Museum**
Mean (standard error)	£2.79 (£0.23)	£3.70 (£0.30)	£3.30 (£0.43)	£4.06 (£0.43)
Median	£1.30	£1.30	£0.00	£0.40

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Sample weighted by museum visitor/gen. pop. weights. Note: All WTP values are calculated as the midpoint interval between the selected payment amount in the payment card and the next highest response on the payment card (except for £0). Summary statistics calculated with inclusion of 'No' at payment principle (coded £0).

The results for **visitor WTP** are within the range of expected values. They are similar to the results of the contingent valuation study for the Natural History Museum, for example, where the mean WTP for an entrance fee was £6.87. Econometric tests support the theoretical validity of the results, i.e. the WTP increases with income and with positive attitudes to culture.

The results for WTP an annual donation as a **non-visitor** to ensure that collections are adequately conserved and maintained, and presented in the best possible way for each institution are also plausible and within the range we would expect for non-user donations. The non-use values are also lower than the range of museum visitor use values, which accords with expectations that users hold higher values for a good or service than non-users (Bateman et al. 2002). They are lower than the estimated non-visitor WTP for the Natural History Museum in our earlier study (£8.29), but arguably this is to be expected given the prominent status that the Natural History Museum has in the eyes of the public. We cannot however, rule out the possibility that order effects

introduced by the survey instrument (see Section 3.5) may also have contributed to lower mean non-use WTP.

The validity tests on non-use WTP are theoretically consistent, with a significant positive association between income and attitudes to culture in the pooled non-visitor model (although we note that this appears to be mainly driven by the non-visitor sample for the Ashmolean museum).

4 Benefit transfer

4.1 Benefit transfer results and validity assessment

In this section, we apply and test the three benefit transfer methods introduced in Section 2.5 (following Brouwer 2000; Johnston et al. 2015). That the mean WTP values across the four museums reported in Section 3 are broadly similar is encouraging in terms of their possible transferability across museum sites. The purpose of this analysis is to evaluate the scope for transferring average WTP values reported later in this study (in Section 4.2) to other existing and/or new museum sites in England.

In the validity tests, we assess whether the estimated WTP values are transferable between study and policy sites, both in terms of the extent of transfer error incurred (as outlined in Equation 9, but also in terms of the statistical significance of the difference between actual and predicted mean WTP (according to the relevant test among those outlined in Table 4.1).

4.1.1 Museum visitor use WTP

The estimated mean WTP for each of the museum sites are reported in the first row of

Table 4.1. Corresponding predictions, based on the pooled study sites, are presented in the second row for each museum when used as a policy site. The greater the % difference between the visitor WTP of the study site and the WTP of the policy site, the greater the transfer error. Both the absolute difference (in absolute monetary terms) and the relative difference (in percentage terms, which is the transfer error) are reported.

Results show that the largest transfer errors (TE) are observed for the World Museum (TE=18%). Conversely, the smallest errors are observed for the NRM (0.3%). The mean difference in absolute levels of WTP between policy and study sites (see hypothesis H1 in Table 2.2) is significant in only one case (World Museum).

In sum, simple unit transfer errors vary between 0% and 18%, safely below what is considered to be an acceptable range (see Section 4.1)

Table 4.1 Museum visitor WTP: Simple unit transfer errors in relative (%) and absolute (£) terms (museum visitor weights applied)

	Policy site			
	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Policy site: Observed average WTP	£7.79	£6.01	£6.86	£7.08
Pooled WTP from study sites (excl. policy site)	£6.56	£7.10	£6.83	£6.81
Transfer error	-15.75%	18.24%	-0.33%	-3.77%
Difference (absolute £)	-£1.23	£1.10	-£0.02	-£0.27
t-test: Difference significant at 5% level?	No	Yes	No	No

Note: Transfer error is calculated according to Equation 9, with $\widehat{WTP}_p$ given by Equation 5.

Inspection of Table 4.2 shows that the adjusted unit approach leads to a slight decrease in transfer errors in the case of two museums (World Museum TE = 12%; NRM TE = 0.25%) and a slight increase in transfer errors in two other cases (Great North Museum TE = 18%; Ashmolean Museum TE = 7%). The range of transfer errors using the adjusted unit transfer approach falls between 0.25% and 18%, which is again within what is considered to be an acceptable range. The mean difference in WTP between policy and study sites (see hypothesis H2 in Table 2.2 is again significant only for the World Museum.

Table 4.2 Museum visitor use WTP: Adjusted unit transfer errors in relative (%) and absolute (£) terms (museum visitor weights applied)

	Policy site			
	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Policy site: Mean income	£40,851.75	£39,807.66	£41,304.15	£45,105.36
Pooled study sites: Mean income	£41,920.43	£41,882.03	£41,271.69	£40,653.97
Policy site: Observed average WTP	£7.79	£6.01	£6.86	£7.08
Pooled WTP from study sites (excl. policy site)	£6.56	£7.10	£6.83	£6.81
Income differential (Policy Income / Study Income)	0.97	0.95	1.00	1.11
Adjusted WTP: Pooled WTP from study sites (excl. policy site), adjusted by Policy/Study income differential	£6.39	£6.75	£6.84	£7.55
Transfer error	-17.90%	12.38%	-0.25%	6.77%
Difference (absolute £)	-£1.39	£0.74	-£0.02	£0.48
t-test: Difference significant at 5% level?	Yes	No	No	No

Note: Transfer error is calculated according to Equation 9, with $\widehat{WTP}_p$ given by Equation 6.

Finally, we consider the function transfer approach. Usually the researcher selects a smaller set of explanatory variables than those presented in the WTP regression models of Table 3.7, opting for variables that are easily available in each site. Here, we specify a simple pooled WTP model with non-parametric mean as the dependent variable[22] and only demographic variables as regressors:

Equation 10

$$WTP = \alpha + \beta_{1Age} + \beta_{2Income} + \beta_{3Distance} + \varepsilon_i$$

where β_{1Age} is calculated from the mid-points of age categories; $\beta_{2Income}$ is a log variable calculated from the mid − points of income categories; and β_{3Dist} is the log of geodesic distance from respondent's postcode to the museum valued (point-to-point distance as the crow flies, in miles). Gender is not included in the reduced model as it is found to be insignificant in the visitor and non-visitor pooled regressions. The simple WTP models are estimated by ordinary least squares, with robust standard errors.

Regression results are presented in Table 4.3. Unsurprisingly, the explanatory power of the reduced regressions is lower than the explanatory power of the best-fit regressions in Table 3.7, given the absence of behavioural and attitudinal determinants. However, the use of reduced regression covariates composed of only those variables which are applicable to all sites is recommended to avoid the problems of over-parameterisation which can lead to increased transfer errors in the transfer function approach (Bateman et al. 2011).

Table 4.3 Museum visitor: Reduced WTP regressions for value transfer (OLS, robust standard errors) (museum visitor weights applied)

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum	Pooled (4 museums)
	b	b	b	b	b
Age, using age midpoint	-0.017	-0.017	-0.031	-0.064*	-0.024*
Log income, using income midpoints	3.144***	2.498***	1.542**	3.025	2.064***
Log geodesic distance to museum	1.482***	0.827***	0.796**	0.582	0.939***
Constant	-30.236***	-22.537***	-11.407	-23.785	-17.671***
Observations	227	256	366	240	1089
Adjusted R-squared	0.125	0.114	0.045	0.067	0.083

*Notes: *** significance at <1%; ** significance at <5%; * significance at <10%. Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the*

[22] Non-parametric mean is calculated from the mid-point intervals of values selected on the payment card.

museum-level ('I do not believe I would have to pay in reality'). Sample weighted by museum visitor weights. Notes: Gross annual household income; averages computed using the midpoints of the income and age categories. Heteroskedasticity-robust standard errors. Note differences in model sample size due to lower omission of observations due to missing data for model covariates.

The transfer errors reported in Table 4.4 are relatively low, varying between 2% in the case of the NRM and 18% in the case of the Great North Museum. This error range again falls below the threshold range proposed in the literature, and is consistent with the transfer errors in using the simple and adjusted transfer approaches above. The mean difference in WTP between policy and study sites (see hypothesis H3 in Table 2.2) is not significant in any of the function transfer tests.

Table 4.4 Museum visitor WTP: Pooled function transfer errors in relative (%) and absolute (£) terms (museum visitor weights applied)

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Policy site: Observed average WTP	£7.79	£6.01	£6.86	£7.08
Study site: Predicted WTP from pooled value transfer function (excl. policy site)	£6.41	£7.01	£7.02	£6.78
Transfer error	-17.68%	16.65%	2.36%	-4.20%
Difference (absolute £)	-£1.38	£1.00	£0.16	-£0.30
t-test: Difference significant at 5% level?	No	No	No	No

Note: Transfer error is calculated according to Equation 9 ***Error! Reference source not found.****, with* $\widehat{WTP}_p$ *given by Equation 8 with regressors as in* Equation 10*.)*

4.1.2 Transfer errors summary: Museum visitor WTP

Table 4.5 Benefit transfer errors (TE) – Museum Use Values - Summary

	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum	Mean TE	Max TE
i) Simple pooled unit transfer	15.8%	18.2%	0.3%	3.8%	9.5%	18.2%
ii) Adjusted for income	17.9%	12.4%	0.2%	6.8%	9.3%	17.9%
iii) Pooled Function transfer	17.7%	16.7%	2.4%	4.2%	10.2%	17.7%

The results displayed in Table 4.5 are indicative that transferring museum visitor use values from the study site (pooled mean WTP) to potential policy sites can be performed with relatively low transfer errors.

In terms of comparison of transfer errors between the three benefit transfer approaches, the maximum observed transfer error across all tests is 18%, which falls below the suggested threshold for transfer errors suggested in the literature (20-40%) (Ready and Navrud 2006). The adjusted unit transfer method performs best overall, yielding the lowest mean transfer error (9%). The maximum transfer error for the adjusted unit transfer is only 0.2pp higher than in the case of function transfer.

Statistical tests of difference between study and policy site WTP are significant in only one of the four tests for both simple and adjusted unit transfer, and in no cases for function transfer. However, the benefit of the adjusted unit transfer over the function transfer approach is that it relies on less data.

Based on these findings, we recommend that policymakers choose which transfer approach to apply depending on data availability and contextual factors, as outlined below.

- The *simple pooled unit transfer method:* Suitable for transferring use WTP values from the four museums we study to policy sites which are sufficiently similar in museum characteristics and visitor demographics.
- Transfer of *benefit functions:* We find that significant factors in the benefit function of use WTP values are the income and age of visitors, and the distance visitors are willing to travel. Where this data is available, policy analysts may prefer to adopt the function transfer approach.
- *Adjusted pooled unit transfer:* this produces similarly low transfer errors and requires less data (only the income differential between study and policy sites), which can be performed through *ex post* adjustment of WTP values.

4.1.3 Museum non-visitor WTP

We report simple pooled unit transfer errors for museum non-visitor WTP values in

Table 4.6.

The results of simple unit transfer of non-use values show that the greatest transfer errors (TE) are observed when predicting the non-use values of the Great North Museum (TE=33%). Conversely, the lowest errors are observed for the NRM (7%). The mean difference in WTP between policy and study sites (see hypothesis H1 in Table **Error! Reference source not found.**2.2) is significant in only one case (Great North Museum).

Simple unit transfer errors vary between 7% and 33%. These are higher than in the visitor WTP transfer tests, but still below the 40% upper-bound suggested by the literature.

Table 4.6 Museum non-visitor non-use WTP: Simple unit transfer errors in relative (%) and absolute (£) terms (general population weights applied)

	Policy site			
	Great North Museum	World Museum	National Railway Museum	Ashmolean Museum
Policy site: Observed average WTP	£2.79	£3.70	£3.30	£4.06
Pooled WTP from study sites (excl. policy site)	£3.71	£3.40	£3.53	£3.26
Transfer error	32.81%	-7.96%	6.88%	-19.70%
Difference (absolute £)	£0.92	-£0.29	£0.23	-£0.80
t-test: Difference significant at 5% level?	Yes	No	No	No

Note: Transfer error is calculated according to Equation 9, with $\widehat{WTP}_p$ given by Equation 8.

In sum, transfer errors obtained for non-use values using the simple unit transfer approach are larger on average (mean TE = 17%) than for the same method as applied to use values (10%, recall Table 4.5). The maximum transfer error is also higher for non-use values (33%) compared with the maximum transfer error for use values using the same transfer method (18%, recall Table 4.5). Non-use WTP is significantly different between the policy and study sites in only one of the four cases.

Based on these results, we conclude that errors are on average higher when transferring non-use values from the pooled study site to non-users at a separate policy site. However, the maximum transfer errors still fall within the 40% threshold deemed acceptable in the literature. These results are solely based on the unit value transfer, however. Mindful of this, our findings for benefit transfer of the non-use values should be taken as indicative only, and we would recommend further research with larger data sets to understand better the transferability of non-use values between non-user populations in the general public.

4.2 Guidance and recommendations for application of values to benefit transfer

In the sections below, we discuss **the relevant considerations and data requirements that should inform the policy analyst's choice of benefit transfer method.** Note that all values are generated from the pooled set of four study sites. The purpose of the tables below is to allow the analyst to select the appropriate benefit transfer method to transfer use and/or non-use values to a fifth policy site.

Note that while, as discussed, within our study it is not possible to reliably test the errors associated with adjusted and function transfer for non-use values within the same population, the policy analyst can still adjust the estimated non-use values based on population differences among non-user populations (for instance, income differentials between national and regional income averages for transfer to specific regional populations). In these circumstances, we provide recommendations on the additional

information that the analyst would need, but with the important caveat that transfer error estimates are not available for these methods.

4.2.1 Pooled simple unit transfer

In our testing, simple unit transfer is found to produce the lowest transfer errors from among the three methods tested on use WTP values, and produces errors within an acceptable range for non-use WTP values too. We outline below the data required and criteria to be taken into account by the policy analyst in applying the simple unit transfer approach (Table 4.7), and the final set of pooled study site values and transfer errors for use in benefit transfer (Table 4.8).

Table 4.7 Simple unit transfer: Data requirements and selection criteria

Data required from study sites	Mean use and non-use WTP
Data required at policy site	Information about the expected characteristics of the policy site, to allow analyst to assess comparability of study and policy sites (following selection criteria, Section 1.3)
When to use this method	When analyst believes there is relative homogeneity in characteristics of the study and policy sites. When data on relevant populations (for function transfer/adjusted transfer) do not exist, or do not vary between study and policy site.
When not to use this method	When study and policy sites differ in site and population characteristics in important ways.

Table 4.8 Use and non-use WTP for benefit transfer: Simple unit transfer

Population	Use/non-use value	Valuation variable	Study site mean WTP (4 sites)	Mean transfer error	Max transfer error
Visitor	**Use**	**Entry fee for access**	£6.42	9.5%	18.2%
Non-visitor	**Non-use**	**Annual donation for conservation, maintenance & presentation of collections**	£3.48	16.8%	32.8%

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). All WTP values are calculated as the midpoint interval between the selected payment amount in the payment card and the next highest response on the payment card (except £0 bids). Summary statistics calculated with inclusion of 'No' at payment principle (coded £0). Pooled visitor sample weighted by visitor weights; pooled non-visitor sample weighted by English nationally representative gender and age group weights.

4.2.2 Pooled adjusted unit transfer

In our testing, *adjusted unit transfer* produced the lowest mean transfer errors in use WTP transfer tests. The adjusted transfer method allows the analyst to adjust the benefits to be transferred from the study to policy sites based on observed differentials between them. Commonly this is based on the average income of the visitor or general population groups associated with the policy institution. The benefit is that this approach is less data-intensive than the transfer function approach and allows for *ex post* adjustment specific to the policy context.

Table 4.9 Adjusted unit transfer: Data requirements and selection criteria

Data required from study sites	Mean use and non-use WTP. Mean annual household income levels
Data required at policy site	Information about the expected characteristics of the policy site, to assess comparability of study and policy sites (following selection criteria, Section 1.3).
When to use this method	When differences are expected between study and policy site populations, it is recommended to adjust transfer for the relevant characteristics (usually income, which is recognised as the strongest theoretical and empirical driver of WTP value).
When not to use this method	When data on income differentials between study and policy site do not exist, or are not significantly different. When study and policy site populations do not differ in average income (e.g. identical national non-user populations)

Error! Not a valid bookmark self-reference. provides the study site mean WTP and income data required by the analyst to perform adjusted unit transfer of use and non-use values to potential policy sites.

Table 4.10 Use and non-use WTP for benefit transfer: Adjusted unit transfer

Population	Use/non-use value	Valuation variable	Study site mean WTP (4 sites)	Study site mean income	Policy site mean income	Income differential (Policy/ Study)	Mean transfer error	Max transfer error
Visitor	**Use**	**Entry fee for access**	£6.42	£40,375	Mean income of visitors to policy site	$\left(\frac{\bar{Y}_p}{\bar{Y}_s}\right)^e$	9.3%	17.9%
Non-visitor	**Non-use**	**Annual donation for conservation, maintenance & presentation of collections**	£3.48	National mean income (ONS statistics)	Mean income of on-user population (e.g. region)		NA	NA

Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). All WTP values are calculated as the midpoint interval between the selected payment amount in the payment card and the next highest response on the payment card (except £0 bids). Summary statistics calculated with inclusion of 'No' at payment principle (coded £0). Pooled visitor sample weighted by visitor weights; pooled non-visitor sample weighted by English nationally representative gender and age group weights.

4.2.3 Pooled benefit function transfer

For benefit transfer of use value, the function transfer approach has acceptable transfer errors. The adjustment of the variables contained in the pooled model may allow for a more robust function transfer model and less error, improving the transfer accuracy. The function transfer approach is also better suited for policy sites which do not share the same characteristics as the current set of study sites, or in situations where this data is unknown. However, the data requirements for the policy site are significant. And there will always remain the large share of unobservable variables which are important for willingness to pay.

For application, the analyst requires data on the socio-demographic and other characteristics of the policy site, specifically age, and annual household income (log), and the average distance travelled by visitors. If data on region of origin exists, it is possible to calculate average distance from the midpoint of this region. These variables are selected because they are commonly available at potential policy sites. Where policy sites involve prospective new museum development, initial scoping and audience prediction data may be used to compare study and policy sites.

Table 4.11 Function transfer: Data requirements and selection criteria

Data required from study sites	Mean WTP. Function coefficients (see Table 4.12)
Data required at policy site	Data on visitors: age, income, distance travelled/region of origin. Where relevant, data on site characteristics: e.g. type of museum, size, etc.
When to use this method	When differences are expected in characteristics of study and policy site characteristics and populations. When policy site visitor demographic data is available, adjustment of the variables contained in the function model may produce more robust function transfer model and less error.
When not to use this method	When the value functions have low explanatory power. When there are few differences between sites to adjust for.

The predicted WTP values and function coefficients required for function transfer are displayed in Table 4.12. We demonstrate how the function transfer is calculated using the museum visitor use WTP regression, below.

Equation 11

$$WTP_{Policy} = -15.944\ (\text{Constant}) - 0.024\ (Age) + 2.064\ (Log\ income) + 0.939(Log\ distance, miles)$$

For the purposes of benefit transfer, the analyst multiplies the regression coefficients from the model by measures of the variables for the policy site to derive the partial WTP

for each of the variables (following Rosenberger and Loomis 2003). For instance, if average age at the policy site is 65, this is multiplied by the coefficient for age in Equation 11 (−0.024) to obtain the partial WTP effect of age at the policy site. The same is repeated for income and distance, summing all of the partial effects to obtain the mean WTP for transfer to the policy site.

Table 4.12 Function transfer: Museum visitor and non-visitor WTP across four study sites: (OLS, robust standard errors) (relevant population weights applied)

	Museum visitor use WTP
Mean transfer error	10.2%
Max transfer error	17.7%
Function transfer values	
Age, using age midpoint	-0.024*
Log income, using income midpoints	2.064***
Log geodesic distance to museum (miles)	0.939***
Constant	-17.671***
Observations	1089
Adjusted R-squared	0.083

*Notes: *** significance at <1%; ** significance at <5%; * significance at <10%. Sample restricted to residents in England aged 16 and over. Sample excludes speedsters (n=65), and inconsistent follow-up responses at the museum-level ('I do not believe I would have to pay in reality'). Sample weighted by museum visitor weights. Notes: Gross annual household income; averages computed using the midpoints of the income and age categories. Heteroskedasticity-robust standard errors. Note differences in model sample size due to lower omission of observations due to missing data for model covariates.*

5 References

Aadland, D., & Caplan, A. J. (2006). Cheap talk reconsidered: New evidence from CVM. *Journal of Economic Behavior & Organization*, *60*(4), 562–578. doi:10.1016/j.jebo.2004.09.006

Arrow, K., & Solow, R. (1993). *Report of the NOAA Panel on Contingent Valuation*. National Oceanic and Atmospheric Administration Washington, DC. http://www.cbe.csueastbay.edu/~alima/courses/4306/articles/NOAA%20on%20contingent%20valuation%201993.pdf. Accessed 28 May 2014

Atkinson, S. E., Crocker, T. D., & Shogren, J. F. (1992). Bayesian exchangeability, benefit transfer, and research efficiency. *Water Resources Research*, *28*(3), 715–722. doi:10.1029/91WR02598

Bakhshi, H., Fujiwara, D., Lawton, R. N., Mourato, S., & Dolan, P. (2015). *Measuring Economic Value in Cultural Institutions* (Cultural Value Project) (p. 103). London, UK: Arts and Humanities Research Council.

Bateman, I., Brouwer, R., Ferrini, S., Schaafsma, M., Barton, D. N., Dubgaard, A., et al. (2011). Making Benefit Transfers Work: Deriving and Testing Principles for Value Transfers for Similar and Dissimilar Sites Using a Case Study of the Non-Market Benefits of Water Quality Improvements Across Europe. *Environmental and Resource Economics*, *50*(3), 365–387. doi:10.1007/s10640-011-9476-8

Bateman, I., Carson, R. T., Day, B., Hanemann, M., Hanley, N., Hett, T., et al. (2002). *Economic Valuation with Stated Preference Techniques: A Manual*. Cheltenham, UK: Edward Elgar.

Bateman, I., & Langford, I. H. (1997). Budget-Constraint, Temporal, and Question-Ordering Effects in Contingent Valuation Studies. *Environment and Planning A*, *29*(7), 1215–1228. doi:10.1068/a291215

Bateman, I., Mace, G. M., Fezzi, C., Atkinson, G., & Turner, K. (2011). Economic Analysis for Ecosystem Service Assessments. *Environmental and Resource Economics*, *48*(2), 177–218. doi:10.1007/s10640-010-9418-x

Bedate, A. M., Herrero, L. C., & Sanz, J. Á. (2009). Economic valuation of a contemporary art museum: correction of hypothetical bias using a certainty question. *Journal of Cultural Economics*, *33*(3), 185–199. doi:10.1007/s10824-009-9098-y

Bergland, O., Magnussen, K., & Navrud, S. (Eds.). (2002). Benefit transfer: testing for accuracy and reliability. In *Comparative Environmental Economic Assessment* (pp. 117–132). Edward Elgar Publishing.

Boyle, K. J., & Bergstrom, J. C. (1992). Benefit transfer studies: Myths, pragmatism, and idealism. *Water Resources Research*, *28*(3), 657–663. doi:10.1029/91WR02591

Boyle, K. J., Kuminoff, N. V., Parmeter, C. F., & Pope, J. C. (2010). The Benefit-Transfer Challenges. *Annual Review of Resource Economics*, *2*(1), 161–182. doi:10.1146/annurev.resource.012809.103933

Brink, P. t., Badura, T., Bassi, S., Daly, E., Dickie, I., Ding, H., et al. (2011). *Estimating the overall economic value of the benefits provided by the Natura 2000 network* (Final Synthesis Report). Brussels, Belgium: European Commission, DG Environment; Institute for European Environmental Policy / GHK / Ecologic.

Brookshire, D. S., & Neill, H. R. (1992). Benefit transfers: conceptual and empirical issues. *Water resources research*, *28*(3), 651–655.

Brouwer, R. (2000). Environmental value transfer: state of the art and future prospects. *Ecological economics*, *32*(1), 137–152.

Brouwer, R., Martin-Ortega, J., Dekker, T., Sardonini, L., Andreu, J., Kontogianni, A., et al. (2015). Improving value transfer through socio-economic adjustments in a multicountry choice experiment of water conservation alternatives. *Australian Journal of Agricultural and Resource Economics*, *59*(3), 458–478. doi:10.1111/1467-8489.12099

Brouwer, R., & Navrud, S. (2015). The Use and Development of Benefit Transfer in Europe. In R. Johnston, J. Rolfe, R. S. Rosenberger, & R. Brouwer (Eds.), *Benefit Transfer of Environmental*

and Resource Values - A Guide for Researchers and Practitioners (pp. 71–84). London, UK: Springer. http://www.springer.com/gb/book/9789401799294. Accessed 26 April 2017

Brouwer, R., & Spaninks, F. A. (1999). The validity of environmental benefits transfer: further empirical testing. *Environmental and resource economics*, *14*(1), 95–117.

Campbell, D., Mørkbak, M. R., & Olsen, S. B. (2017). Response time in online stated choice experiments: the non-triviality of identifying fast and slow respondents. *Journal of Environmental Economics and Policy*, *6*(1), 17–35. doi:10.1080/21606544.2016.1167632

Carlsson, F., Frykblom, P., & Johan Lagerkvist, C. (2005). Using cheap talk as a test of validity in choice experiments. *Economics Letters*, *89*(2), 147–152. doi:10.1016/j.econlet.2005.03.010

Carlsson, F., Kataria, M., Krupnick, A., Lampi, E., Löfgren, Å., Qin, P., & Sterner, T. (2013). The truth, the whole truth, and nothing but the truth—A multiple country test of an oath script. *Journal of Economic Behavior & Organization*, *89*, 105–121. doi:10.1016/j.jebo.2013.02.003

Carlsson, F., & Martinsson, P. (2006). Do Experience and Cheap Talk Influence Willingness to Pay in an Open-Ended Contingent Valuation Survey? In *Working Papers in Economics 109*.

Carson, R. T. (2012). Contingent valuation: a practical alternative when prices aren't available. *Journal of Economic Perspectives*, *26*(4), 27–42. doi:10.1257/jep.26.4.27

Champ, P. A., & Bishop, R. C. (2001). Donation payment mechanisms and contingent valuation: an empirical study of hypothetical bias. *Environmental and Resource Economics*, *19*(4), 383–402.

Crossick, G., & Kaszynska, P. (2016). *Understanding the value of arts & culture. The AHRC Cultural Value Project*. London, UK: AHRC.

Cummings, R. G., & Taylor, L. O. (1999a). Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method. *The American Economic Review*, *89*(3), 649–665.

Cummings, R. G., & Taylor, L. O. (1999b). Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method. *The American Economic Review*, *89*(3), 649–665.

Darling, A. C., & et al. (2000). Uncertainty in cost-benefit analysis based on referendum contingent valuation. *mpact Assessment and Project Appraisal, 18*(2), 125–137. doi:10.3152/147154600781767466

Desvouges, W. H., Naughton, M. C., & Parsons, G. R. (1992a). Benefits transfer: conceptual problems in estimating water quality benefits using existing studies, *28*(3), 675–683.

Desvouges, W. H., Naughton, M. C., & Parsons, G. R. (1992b). Benefits transfer: conceptual problems in estimating water quality benefits using existing studies, *28*(3), 675–683.

Desvousges, W. H., Johnson, F. R., Banzhaf, H. S., & others. (1998). *Environmental policy analysis with limited information: principles and applications of the transfer method.* Edward Elgar Publishing. http://www.cabdirect.org/abstracts/19991801535.html. Accessed 21 November 2016

Desvousges, W. H., Naughton, M. C., & Parsons, G. R. (1992). Benefit transfer: Conceptual problems in estimating water quality benefits using existing studies. *Water Resources Research*, *28*(3), 675–683. doi:10.1029/91WR02592

Eftec. (2000). *Valuing Our Recorded Heritage: Final Report*. London, UK: Economics for the Environment Consultancy.

Eftec. (2005). *Valuation of the Historic Environment The scope for using results of valuation studies in the appraisal and assessment of heritage-related projects and programmes* (p. 90). London, UK: Economics for the Environment Consultancy. http://www.english-heritage.org.uk/publications/valuation-historic-environment/valuation-historic-environment-final-rep.pdf

Eftec. (2009). *Valuing Environmental Impacts: Practical Guidelines for the Use of Value Transfer in Policy and Project Appraisal*. London, UK: Department for Environment, Food and Rural Affairs. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/182376/vt-guidelines.pdf.

Hausman, J. (2012). Contingent valuation: from dubious to hopeless. *Journal of Economic Perspectives*, *26*(4), 43–56. doi:10.1257/jep.26.4.43

HM Treasury. (2011). *The Green Book: Appraisal and Evaluation in Central Government* (pp. 1–114). HM Treasury. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220541/green_book_complete.pdf. Accessed 21 May 2014

Jacquemet, N., Joule, R.-V., Luchini, S., & Shogren, J. F. (2013). Preference elicitation under oath. *Journal of Environmental Economics and Management*, *65*(1), 110–132. doi:10.1016/j.jeem.2012.05.004

Johnston, R., Rolfe, J., Rosenberger, R. S., & Brouwer, R. (2015). *Benefit Transfer of Environmental and Resource Values - A Guide for Researchers and Practitioners*. London, UK: Springer. http://www.springer.com/gb/book/9789401799294. Accessed 26 April 2017

Johnston, R., & Rosenberger, R. S. (2010). Methods, trends and controversies in contemporary benefit transfer. *Journal of Economic Surveys*, *24*(3), 479–510.

Kahneman, D., & Knetsch, J. L. (1992). Valuing public goods: The purchase of moral satisfaction. *Journal of Environmental Economics and Management*, *22*(1), 57–70. doi:10.1016/0095-0696(92)90019-S

Kaul, S., Boyle, K. J., Kuminoff, N. V., Parmeter, C. F., & Pope, J. C. (2013). What can we learn from benefit transfer errors? Evidence from 20 years of research on convergent validity. *Journal of Environmental Economics and Management*, *66*(1), 90–104. doi:10.1016/j.jeem.2013.03.001

Kristofersson, D., & Navrud, S. (2007). Can use and Non-use Values Be Transferred Across Countries? In S. Navrud & R. Ready (Eds.), *Environmental Value Transfer: Issues and Methods* (pp. 207–225). Springer Netherlands. http://link.springer.com/chapter/10.1007/1-4020-5405-X_11. Accessed 25 April 2017

Landry, C. E., & List, J. A. (2007). Using Ex Ante Approaches to Obtain Credible Signals for Value in Contingent Markets: Evidence from the Field. *American Journal of Agricultural Economics*, *89*(2), 420–429. doi:10.1111/j.1467-8276.2007.01017.x

List, J. A., & Lucking-Reiley, D. (2000). Demand reduction in multiunit auctions: Evidence from a sportscard field experiment. *American Economic Review*, 961–972.

Loomis, J. (1992a). The evolution of a more rigorous approach to benefit transfer: Benefit function transfer. *Water Resources Research*, *28*(3), 701–705. doi:10.1029/91WR02596

Loomis, J. (1992b). The evolution of a more rigorous approach to benefit transfer: benefit function transfer. *Water Resources Research*, *28*(3), 701–705.

Loomis, J. (2015). The Use of Benefit Transfer in the US. In R. Johnston, J. Rolfe, R. S. Rosenberger, & R. Brouwer (Eds.), *Benefit Transfer of Environmental and Resource Values - A Guide for Researchers and Practitioners* (pp. 61–70). London, UK: Springer. http://www.springer.com/gb/book/9789401799294. Accessed 26 April 2017

Lusk, J. L. (2003). Effects of Cheap Talk on Consumer Willingness-to-Pay for Golden Rice. *American Journal of Agricultural Economics*, *85*(4), 840–856. doi:10.1111/1467-8276.00492

Maddison, D., & Foster, T. (2003). Valuing congestion costs in the British Museum. *Oxford Economic Papers*, *55*(1), 173–190. doi:10.1093/oep/55.1.173

Maddison, D., & Mourato, S. (2001). Valuing different road options for Stonehenge. *Conservation and Management of Archaeological Sites*, *4*(4), 203–212. doi:10.1179/135050301793138182

Mahieu, P.-A., Riera, P., & Giergiczny, M. (2012). The influence of cheap talk on willingness-to-pay ranges: some empirical evidence from a contingent valuation study. *Journal of Environmental Planning and Management*, *55*(6), 753–763. doi:10.1080/09640568.2011.626524

Malhotra, N. (2008). Completion Time and Response Order Effects in Web Surveys. *Public Opinion Quarterly*, *72*(5), 914–934. doi:10.1093/poq/nfn050

Mourato, S., Fimereli, E., Contu, D., Gaskell, C., & Boniatti-Pavese, C. (2014). *The Economic Benefits of Cultural Built Heritage Interiors Conservation from Climate Change Damages in Europe*. Grantham Research Institute on Climate Change and the Environment.

Mourato, S., & Mazzanti, M. (2002). Economic valuation of cultural heritage: evidence and prospects.

Navrud, S. (2002). *The State-Of-The-Art on Economic Valuation of Noise* (Final Report to European Commission DG Environment) (p. 12). Brussels, Belgium: Department of Economics and Social Sciences Agricultural University of Norway.

Navrud, S., & Ready, R. (2007). *Environmental Value Transfer: Issues and Methods*. Dordrecht, Netherlands: Springer. http://www.springer.com/us/book/9781402040818. Accessed 21 November 2016

Noonan, D. S. (2003). Contingent Valuation and Cultural Resources: A Meta-Analytic Review of the Literature. *Journal of Cultural Economics*, *27*(3–4), 159–176. doi:10.1023/A:1026371110799

OECD. (2014). *The Cost of Air Pollution: Health Impacts of Road Transport*. London, UK: OECD Publishing. http://dx.doi.org/10.1787/9789264210448-en

Pearce, D., & O'zdemiroglu, E. (2002). *Economic valuation with stated preference techniques*. London, UK: Department for Transport, Local Government and the Regions. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/191522/Economic_valuation_with_stated_preference_techniques.pdf

Powe, N. A., & Bateman, I. (2003). Ordering effects in nested "top-down" and "bottom-up" contingent valuation designs. *Ecological Economics*, *45*(2), 255–270. doi:10.1016/S0921-8009(03)00084-3

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.

Pung, C., Clarke, A., & Patten, L. (2004). Measuring the economic impact of the British library. *New Review of Academic Librarianship*, *10*(1), 79–102. doi:10.1080/13614530412331296826

Ready, R., & Navrud, S. ale. (2006). International benefit transfer: Methods and validity tests. *Ecological economics*, *60*(2), 429–434.

Rosenberger, R., & Loomis, J. (2003). Benefit transfer. In *A primer on nonmarket valuation* (pp. 445–482). Springer. http://link.springer.com/chapter/10.1007/978-94-007-0826-6_12. Accessed 21 November 2016

Rosenberger, R. S., & Loomis, J. B. (2003). Benefit transfer. In *A primer on nonmarket valuation* (pp. 445–482). Springer. http://link.springer.com/chapter/10.1007/978-94-007-0826-6_12. Accessed 21 November 2016

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage learning.

Smith, V. K. (1992). On separating defensible benefit transfers from "smoke and mirrors." *Water Resources Research*, *28*(3), 685–694. doi:10.1029/91WR02594

Walsh, R. G., Johnson, D. M., & McKean, J. R. (1992). Benefit transfer of outdoor recreation demand studies, 1968–1988. *Water Resources Research*, *28*(3), 707–713. doi:10.1029/91WR02597

Whitehead, J. C., Morgan, O. A., & Huth, W. L. (2015). Benefit transfers with the contingent valuation method. In *Benefit Transfer of Environmental and Resource Values - A Guide for Researchers and Practitioners* (pp. 119–140). London, UK: Springer. http://www.springer.com/gb/book/9789401799294. Accessed 26 April 2017