

# Measuring Economic Value in Cultural Institutions

A report commissioned by the Arts and Humanities  
Research Council's Cultural Value Project



**Hasan Bakhshi, Daniel Fujiwara, Ricky Lawton,  
Susana Mourato and Paul Dolan**

Corresponding author: [d.fujiwara@simetrica.co.uk](mailto:d.fujiwara@simetrica.co.uk)

## **Executive summary**

### **1. Background**

There are few more deeply contested relationships in cultural policy than that between economics and cultural value. The way that we measure the value of cultural institutions matters. They are valued in distinct ways by those who use it and those that do not, as well as by different groups in society. Our choice of measurement affects whose values we capture. For this reason, arts funders have recently advocated a 'holistic' assessment of the benefits of culture, which goes beyond the economic and cultural to encompass the wellbeing, societal and educational value of culture (Arts Council England, 2014).

This broader definition of value aligns with the concept of value used in cost-benefit analysis (CBA), the standard evaluation methodology used in UK policymaking. Specifically, CBA is the recommended approach in HM Treasury's Green Book Guidance and it also forms the core evaluation methodology in most other OECD countries and in many international organisations (Organisation for Economic Cooperation and Development, 2006; US Office of Management and Budget, 1992). CBA assesses the costs of an action, intervention or investment against the benefits that it creates for society. The benefits should account for all of the possible channels through which society is benefited, and hence will include many of the aspects of value discussed by Arts Council England (2014). CBA requires that benefits and costs be measured in the same metric to make them comparable, which is usually in monetary terms (although versions of CBA can be carried out with non-monetary outcome measures). For cultural goods and services, this means the valuation of intangible and often hard-to-quantify outcomes. Monetary valuation methods are also used extensively in other forms of policy evaluation and impact assessment such as social return on investment (SROI) and social accounting.

Robust valuation of cultural goods and services, therefore, has a number of key benefits. First, it allows cultural institutions to demonstrate in quantitative terms the value that they create for society in a manner that is consistent with best-practice methodology within government. Second, it permits policymakers to assess interventions and investment decisions within the cultural sector using accepted forms of policy evaluation such as CBA. Ultimately, one would expect this to lead to a better and more efficient use of resources in the sector. And third, since – as we shall discuss in this paper – the notion of economic value is inextricably linked to an individual's wellbeing and quality of life, valuation of cultural goods and services provides information on the extent to which people's lives are improved through culture, which will be of key policy and general interest aside from the implications that this might have for CBA and other methods.

Standard economic impact studies have generally tended to value the market benefits of culture, focusing on impacts on jobs, value added and exports. It is harder to assess the value of culture to individuals in society, since in many cases market prices do not exist, and cultural institutions are free at the point of use. But an understanding of this value is required for the main accepted forms of assessment and policy evaluation. And failing to adequately value cultural benefits risks an underappreciation of the social value of cultural investments. The challenge is, therefore, to capture the non-market benefits of culture in a way that takes into account the range of values it offers both to 'users' and the public as a whole.

Although common in areas such as transport or the environment, practical applications of economic valuation techniques in the UK's cultural sector remain few and far between, and where they have been conducted the results are not generally well known. As a result, policy debates in this area have been fragmented and curiously ungrounded in empirical evidence. It should be acknowledged however that a large number of valuation methods exist and that it is often hard to find consensus in terms of which approach is appropriate and this is especially the case in cultural sector valuation.

### **2. Research objectives of the study**

Our research aims to address the evidence gap by applying two broad classes of valuation techniques – **stated preference methods** and the **wellbeing valuation approach** – to the work of two of the UK's premier cultural institutions, the **Natural History Museum (NHM)** and **Tate Liverpool (TL)**, and drawing conclusions on the strengths and weaknesses of these methodologies as applied to culture.<sup>1</sup> It is important to note that our aim here is not to come up with a total value for these institutions as such, but to uncover the value of some of the key non-market services they provide. Moreover, our study is not designed to compare the values derived across the two institutions and hence the values obtained are not comparable. Firstly, the NHM and TL are two very different institutions offering a different array of cultural services and our study focuses on

<sup>1</sup> This report was co-funded by the AHRC's Cultural Value Project and the Department for Culture, Media and Sport.

different subsets of these services in each case. Secondly, the valuation methods used for the two institutions are different in many respects. As we will discuss, even within the contingent valuation survey, for example, different scenarios, value elicitation methods and payment vehicles are used, and different samples of the UK population are surveyed.

A comparative analysis of stated preference and wellbeing valuation methods is important for a number of reasons. First, the cultural sector has generally been hesitant towards traditional economic approaches to valuation. There are valid questions about whether monetary values should ever be applied to culture. However, as discussed, the benefit of attaching monetary value to non-market goods, like culture, is that it increases the likelihood that these values will be considered in economic decision-making. Both stated preference and wellbeing valuation are endorsed by HM Treasury's Green Book on cost-benefit analysis. A detailed comparison of the methods therefore holds out the promise of a more fit-for-purpose approach to economic valuation in the cultural sector that is also recognised by policymakers.

Second, the research is the first comparison of stated preference and wellbeing valuation methods in the cultural sector and amongst the first ever conducted in any sector. It also provides the first monetary valuation estimates of access to and services provided by the NHM and TL.

Finally, the study offers the first comprehensive subjective wellbeing investigation and valuation of NHM and TL visitation using a number of complementary wellbeing approaches.

As such, the report has two main audiences: **funders** charged with evaluating the case for individual cultural investments, and **cultural institutions** who are putting these cases together.

Our **main research questions** are as follows:

- *Can stated preference-based and wellbeing valuation methods deliver convincing and robust valuations in cultural institutions?*
- *How consistent or conflicting are the valuations derived from applying these two methods?*
- *Can techniques from the two valuation methods be combined to optimise their use by cultural institutions and funders?*

This paper presents answers to these research questions using data from bespoke on-site visitor surveys and online general population surveys collected with the assistance of the Natural History Museum in October and November of 2014 and of Tate Liverpool from February to April of 2015.

The focus of this study is on the **primary benefits** of cultural institutions, though we acknowledge they are only one dimension – albeit a key one – of its overall value. That is, we value the **direct** impact of engagement with the cultural institution under study and of the services it provides on the public's wellbeing. The Total Economic Value (TEV) of a cultural institution is divided between **use value** and **non-use value**. Use values include the *direct use* benefits to visitors of cultural institutions, for instance recreational, leisure, and entertainment activities, as well as education, inspiration and knowledge. *Indirect use* benefits could arise in the form of enhanced community image, and social interaction. *Option value* refers to benefits from a potential future use of the institution and its services. However, cultural services can be valued even by those who do not directly use them. Non-use values can come from simply knowing that others will benefit, either now, or in the future. Non-use value may also be derived from simply knowing that the cultural institution exists. For example, this may refer to a sense of pride associated with the existence of a renowned UK institution. In our study we capture use values as well as elements of non-use value.

Primary benefits differ from another type of benefit known as **secondary benefits**. Secondary benefits refer to induced benefits to society more widely rather than for an individual directly. These are often measured in terms of impacts on public funds, resources and services. For example, if a cultural activity or programme led to improved health outcomes for an individual this would have a direct primary benefit for the individual's quality of life and wellbeing, which will then lead to an induced secondary benefit in terms of possible reductions in healthcare expenditure. Secondary benefits like these play an important role in policy evaluation. However, this type of impact analysis is a separate area of evaluation that requires a very different set of methodologies and is therefore out of the scope of the present study.

### 3. The valuation methods

The study undertakes an explicit **comparison between two methods of economic valuation** using a popular stated preference method, **contingent valuation (CV)**, and **wellbeing valuation (WV)**. At the heart of any economic valuation method is the concept of wellbeing or welfare. Economic value is the equivalent or compensating amount of money that would generate the same effect on an individual's wellbeing as the non-market good, here cultural engagement. Monetary value in

economic terms is, therefore, inextricably linked to welfare; indeed, it is a direct measure of changes in an individual's wellbeing. CV and WV methods primarily differ in the way that they define wellbeing. CV uses a preference account of wellbeing whereby it is assumed that "what would be best for someone is what would best fulfil his desires" (Parfit, 1984, p. 4). Under this account of welfare, CV methods present a hypothetical situation where the non-market good (in this case some element of the cultural institution under study) is presented to visitors or the general public, and the respondent is then asked to state their maximum willingness to pay (WTP) in monetary terms. In the case studies developed at NHM and TL, monetary value is elicited through a hypothetical entry fee or a donation. Using best-practice techniques, this can provide estimates of economic value.

An alternative to the CV technique is the WV approach, which takes people's self-reports of their own wellbeing as a measure of their welfare. The method looks at how experiencing certain outcomes, such as regularly visiting a cultural institution, impacts on people's quality of life through subjective wellbeing (SWB) measures. This measure of SWB could be, for example, life satisfaction or how happy one feels. An economic value is ascertained by estimating the amount of money that would generate the equivalent impact on their wellbeing (e.g. life satisfaction) as engagement with culture, under the empirically supported assumption that income has a positive effect on measures of SWB.

Contingent valuation and wellbeing valuation provide two distinct ways of estimating economic value based on two substantively different approaches to measuring human welfare: preferences and subjective wellbeing. As such, we would not expect them to give the same valuation results – for example, it may be that people desire things that do not impact positively on their SWB, or people may care about things for more than just the impact that they have on their SWB. Our study assesses the extent to which the valuation results from these two methods are comparable; whether there are any clear advantages of one approach over another in the specific context of valuing the work of a cultural institution; and whether it is desirable to combine the methods in a 'hybrid' approach.

### 3.1. A comparison of the valuation methods

There are a number of comparative strengths and weaknesses of the CV and WV approaches. CV is a well-established non-market valuation technique that has been developed and improved since the 1980's. The key methodological advantage of CV over other non-market valuation approaches is that it can measure values and benefits that would not be revealed under market conditions, such as *non-use* values. Importantly, it also offers the opportunity to measure benefits associated with changes that have not yet happened, that is, *future* changes (such as when a cultural institution undergoes a programme of conservation work, or invests to avoid an irreversible risk to its work). However, CV is subject to a number of biases, in particular hypothetical bias, where individuals' stated WTP may be significantly larger than actual WTP due to the hypothetical nature of the survey (Arrow and Solow, 1993). People may focus only on the particular aspects of the outcome described in the CV survey, and this may not reflect in any way how people would actually experience these conditions or states in real life (Hausman, 2012). Problems also arise with CV when people have insufficient information to make informed choices, and a growing literature shows that preferences can in this case be influenced by contextual factors. A number of best practice methods are available to address such bias problems, and we test their suitability to the cultural sector in the case studies included here.

Last, but not least, some have raised ethical concerns about the appropriateness of using WTP at all to value services like health and culture. This may, for example, be because value is related to ability to pay and the prevailing income distribution may be seen as inequitable; or because using money to value health and culture may send an undesirable signal (namely that health and cultural services are just like any other commodity bought and sold in the market place) (Fujiwara and Dolan, 2014).

WV assesses the importance of a factor in someone's life without having to ask them to attribute a value to that factor. The WV approach is based on real experiences and not – as in stated preference approaches – on how people *imagine* they might be affected by a change. In wellbeing analysis, it is possible to look at people's actual experiences of engaging with a cultural institution when they are living life as they normally do, and attach values to these conditions. This eliminates so-called focusing illusion problems found in CV. WV also avoids problems related to hypothetical bias and contextual influences because WTP is not elicited. And finally, because people are not asked to place a value on the non-market good directly, the WV approach avoids ethical difficulties which arise when respondents are asked to assess the value of a cultural good or service in monetary terms. For many this may be an ethically difficult or even objectionable task to undertake.

However, there are concerns that SWB measures may not be sensitive to small differences in actual wellbeing (Johns and Ormerod, 2007). This may be of particular concern with impacts from museum or gallery visits or other discrete activities where the wellbeing impact might be expected to be limited. WV cannot be used to measure non-use values in any obvious

way, since the individual needs to have experienced ('used') the good or service for it to show up in their wellbeing responses. Peoples' evaluation of their own wellbeing may suffer from imperfect recollection of past experiences, or be influenced by contextual factors present at the time of survey (Kahneman and Krueger, 2006).

In sum, there are a number of comparative advantages and disadvantages associated with CV and WV approaches which precludes any prior judgement on which works 'best' when applied in the cultural sector. Hence the focus in this paper is on an empirical examination of the issues in the context of two practical case studies, of the Natural History Museum and Tate Liverpool.

## 4. Results

### 4.1. Different measures of value

In the CV and WV studies, we estimate a range of values and care must be taken when using and interpreting these values. Specifically, we estimate three types of values with different meanings depending on the context they refer to.

**(i) Visitor use value** provides an estimate of direct use value from the perspective of the current visitors to the cultural institution. This is elicited as an entry fee in the case of the NHM, and as a donation in the case of TL. As discussed, this is one reason why the use values are not directly comparable between the two case studies. Generally speaking, visitor use value is comparable with the estimates that are derived from the WV approach and the hybrid CV/WV method we explore which assesses the compensation values for visitors for a hypothetical closure of each institution to visitors.

**(ii) Non-use value among visitors** represents visitors' WTP for elements of the cultural institution that they do not directly benefit from, such as its conservation and research work (in the case of the NHM) or its community outreach and other work outside the gallery (in the case of TL). In both cases, visitor non-use value can be interpreted as additional to visitor use value. Although both values refer to a one-off visitor donation, it is not possible to combine visitor non-use values between the two case studies, as they refer to very different, specific types of work that each institution performs outside the museum or gallery, and are elicited via different types of donations in each case (a voluntary top-up to an entry fee in the case of the NHM, and a separate donation in the case of TL).

**(iii) Non-use/option value among the general population** is elicited for the institution as a whole, including the work it does both inside and outside the museum/gallery. It is not, however, possible to equate non-use value amongst visitors to non-use value in the general population, since they refer to different elements of the cultural institution being valued, and have been elicited using different payment mechanisms (annual donations vs. one-off donations) and from different population samples.

The values elicited across the two case studies therefore refer to different scenarios, payment settings, and populations. 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, and among visitors whether we are interested in direct use or non-use values, as well as the influence that the payment mechanism (entry fee, top-up donation, one-off or annual donation) has on the size of the value elicited. It is therefore important that each value is used in the correct policy context.

### 4.2. Contingent valuation values

We find that CV is a viable approach for measuring economic values in the context of cultural institutions. The method produces realistic values that vary in ways that are consistent with economic theory and previous findings, across different institutions, scenarios, payment vehicles and population groups.

Specifically, CV gives visitor use values that are of a plausible magnitude, and in line with prices currently charged for paid exhibitions in cultural institutions in the UK. This is the case whether the payment mechanism is paid entry (NHM) or a donation to support the institution's work (TL). As expected, the amounts individuals are willing to pay (WTP) or donate are strongly affected by income. Positive attitudes towards the institutions drive WTP in ways that we would expect to be important on theoretical grounds also matter.

Contingent valuation also gives plausible estimates of non-use values – elicited in both institutions using a willingness to donate mechanism – associated with conservation and research work in the case of the NHM, and community and outreach work in the case of TL. These values also vary in ways predicted by theory, with income being a key determinant.

However, in the visitors' non-use value scenario in both cases, agreement with the statement that (temporary) closure of the institution in question would reduce life satisfaction is a significant positive driver of willingness to donate, suggesting that, despite best efforts, it is not possible to fully disentangle use and non-use value in the responses of visitors as they are likely to be intrinsically bundled together. Consistent with this, in the NHM visitors' survey, time spent in the museum turns out to be a strong driver of hypothetical donations for the museum's conservation and research work.

In the online general population survey, stated non-use/option values also accord strongly with theoretically relevant factors for both institutions. In both the NHM and TL, income levels and agreement with the statement that the institution provides value even for those who do not use it – a direct indicator of non-use value – are significant drivers of willingness to donate. Given their different contexts, in some cases the WTP determinants differ in the two case studies. At the NHM, likelihood of a future visit is a significant positive driver of WTP, as we are also capturing option values in the online survey. Familiarity with the NHM in Tring is also a strong predictor of higher willingness to donate, which is consistent with the assumption that placing a non-use value on a cultural institution like the NHM is contingent on one's knowledge of its existence and awareness of the value of its work. Education levels are a significant and positive driver of non-use value too, consistent with previous CV studies of cultural institutions. In the case of TL, familiarity with its community and outreach work is also a strong predictor of willingness to donate. We also find that wider engagement with cultural institutions, as represented by whether the respondent has recently attended other galleries, is positively significant in this case.

### 4.3. Wellbeing valuation values

Our analysis of the SWB data found that in both institutions there is a strong positive association between activities at the institution and momentary wellbeing indicators, measured as how happy people feel and their sense of purpose. This is after controlling for a range of other factors that impact on momentary wellbeing. In the NHM attending paid exhibitions is associated with a particularly high sense of purpose and happiness. In TL, attending lectures, talks or workshops is associated with a high sense of purpose.

These results are in line with previous studies using similar data that show that cultural activities generally rank highly in terms of momentary measures of SWB (Fujiwara et al. 2014). They provide an independent rationale and reason to explain why people value and are willing to pay to visit cultural institutions like the NHM and TL, and we would therefore expect there to be an association between people's engagement with these institutions and SWB over the longer term.

The analysis of the life satisfaction data, however, does not find the positive association between Tate Liverpool visits over the past 12 months and life satisfaction to be statistically significant. This finding could conceivably reflect a number of factors other than the absence of a genuine effect, including measurement error (e.g. people misreporting the number of times they have actually visited TL), and the possibility that the sample size is insufficient to allow a significant impact to be detected.

In the case of the NHM, visits over the past 12 months are associated with very high levels of life satisfaction and hence very high monetary values when measured using wellbeing valuation. The magnitude of this effect is not supported by other studies of cultural activities and wellbeing, and is nearly as large as some major life events such as employment. Selection bias may always be an issue in life satisfaction models, but we have controlled for a much richer array of selection factors in this analysis than previous studies of cultural activities and wellbeing – which typically use secondary data – have been able to do. We speculate that the estimated large life satisfaction impact is driven by an artefact of the survey, namely a priming effect whereby people who care for or like the NHM are being reminded of the importance of the NHM in the survey, which artificially inflates their wellbeing at the time of the survey (along the same lines as previous studies that have reported impacts of trivial contextual factors on life satisfaction).

The problems that we have encountered may conceivably be due to a combination of restricted sample sizes and priming effects, issues that the national surveys do not suffer from. Equally, it may be that the life satisfaction impacts of engaging with a single cultural institution will always be too small to detect robustly. We tentatively conclude therefore that the WV method is more suitable for valuing engagement with culture over time and more generally than for discrete experiences with individual institutions. The conventional approach to WV is best suited to the valuation of non-marginal changes in states or conditions or for measuring the value of frequent activities where priming effects are not an issue. These are precisely the instances to which CV is not so well suited. Cultural institutions seeking to carry out surveys on wellbeing with visitors and members should note the potential for problems like priming effects.

Although, as noted, we control for a range of other important determinants of wellbeing we acknowledge that it may be difficult to ascertain *causal* relationships in the data using the standard wellbeing approach (it may just be that certain types of people with higher wellbeing anyway visit these institutions). To anticipate this, we also test a novel '**vignette study**' approach in WV. This method presents survey respondents with a hypothetical scenario, in which they are asked to report their levels of SWB if they were able to visit the institutions more frequently (keeping all else in their life constant). This is designed to elicit the impact on life satisfaction of higher frequency visits whilst controlling for other factors. In the NHM, life satisfaction is responsive to visit

frequency in the way expected, but the estimated value per visit is around 6 times higher than the WTP and WTA equivalent values. In TL, the life satisfaction scores do not increase with higher visit frequency in the vignette scenarios. Although vignettes like these have been used in other fields of research, these results raise questions about the current validity of using vignettes to assess changes in wellbeing in cultural settings (clearly though, as a first such study, further research is needed).

#### 4.4. Hybrid approach combining contingent valuation and wellbeing valuation

We test a novel '**hybrid contingent valuation/wellbeing valuation**' approach, where we ask respondents directly how much monetary compensation they would require if they were not able to visit the institution for one year due to a hypothetical closure of the institution such that their life satisfaction would remain unaffected. Crucially, compensation is only offered to those who previously indicated that their life satisfaction would decrease if the institution were temporarily closed. The hybrid contingent/wellbeing valuation approach, which takes the hypothetical setting of CV and combines it with the underlying theory of WV, offers an approach to valuing cultural services on a per visit basis without the need to set out a WTP scenario based on an entrance fee, which can often be problematic (and possibly send out politically unhelpful messages) in a cultural setting. The hybrid approach also seemingly has the added advantage of reducing the well-known disparities between WTA and WTP valuations that are often found in the literature (something we also find, though to a lesser extent, in the NHM study).

The 'willingness to accept' (WTA) scenario in the hybrid approach works well and, notably, gives values per visit that are comparable to the equivalent WTP values. This method also shows that people do feel that visits to the institutions impact on their life satisfaction over the longer term, suggesting that we should not perhaps read too much into the results of the statistical analysis of life satisfaction on recent visits mentioned earlier.

Although WTP has now become the preferred monetary elicitation method in the CV literature, it is acknowledged that there are times when WTA is warranted, for example when property rights are such that respondents believe they have some *intrinsic* right to the good or service in question (culture is arguably a good example of such a case). In instances like these we have shown that a hybrid approach grounded also in the theory of WV can deliver plausible WTA values.

#### 4.5. Summary: main results

We provide summary results of the contingent valuation and wellbeing valuation performed on the Natural History Museum and Tate Liverpool.

Visitor use values are on average £6.65 for the NHM and £10.83 for TL. These figures are of a plausible magnitude compared to prices currently charged for paid exhibitions in UK museums. Visitor WTP for Tate Liverpool was elicited via a donation rather than entrance fee payment mechanism, meaning that the two values are not directly comparable.

Average visitor non-use value in the NHM study is £2.78; it represents visitors' WTP an additional donation on top of an entry-fee to support the research and conservation work of the NHM. This amount is consistent with our expectations that a top-up donation would be of a lower order of magnitude than WTP an entry fee. Visitor non-use value of the work of TL in the wider community, elicited as a donation, averages £8.00.

The online survey captures non-use and option values for the general UK population (excluding N. Ireland) as an annual donation. In the NHM study the online survey values the research and conservation work of the NHM, while the TL study values the work of TL inside and outside the gallery. Values are not widely divergent (£8.29 and £6.10 respectively). In both cases this value is conceptually distinct to both use and non-use values elicited in the visitor survey, due to differences in the scenario and payment vehicles used.

What we call the hybrid contingent-wellbeing valuation approach, i.e. the WTA scenario based on life satisfaction, for a scenario involving a hypothetical closure of the NHM or TL for one year, provides plausible values per visit of £6.89 and £7.13 respectively. As discussed above, the vignette per visit life satisfaction value does not provide theoretically consistent findings.

**Table 1. NHM and TL valuation results**

Survey	Use/non-use value	Valuation variable	Mean	Median
<b>Natural History Museum</b>				
Visitor	Use	Combined WTP for access to the NHM (entrance fee)	£6.65	£6.00
Visitor	Non-use (visitors)	Visitor donation (WTP)	£2.78	£2.25
Online	Non-use and option (general population)	Annual donation (WTP)	£8.29	£4.25
Visitor and online		Compensation for closure (per visit) (WTA)	£6.89	N/A
Visitor and online		Vignette wellbeing value (per visit)	£40.42	N/A
<b>Tate Liverpool</b>				
Visitor	Use	Visitor donation for work inside the gallery (WTP)	£10.83	£5.50
Visitor	Non-use (visitors)	Visitor donation for work outside the gallery (WTP)	£8.00	£2.50
Online	Non-use and option (general population)	Annual donation for work inside and outside the gallery (WTP)	£6.10	£0
Visitor and online		Compensation for closure (per visit) (WTA)	£7.13	N/A
Visitor and online		Vignette wellbeing value (per visit)	N/A	N/A

Note: values calculated using mid-point of intervals from the payment cards. Standard WV values are excluded due to the discussed problems concerning priming effects in the survey.

#### 4.6. Main conclusions

Based on our findings, our first main conclusion is that contingent valuation performs well in estimating credible, intuitive and theoretically-consistent use and non-use values in different types of cultural institution, using a variety of scenarios, payment vehicles and population groups. In addition, when undertaking field work we employ a number of best practice techniques which we suggest should be seen as a requirement for contingent valuation as applied to cultural institutions going forwards. These include the use of:

- Detailed descriptions of the good/service, including photographs;
- Precise descriptions of the payment vehicle;
- Pre-testing to check the survey's ease of use;
- Application of the most appropriate payment mechanisms to the circumstances of each site (e.g. entrance fees or voluntary donations) to elicit direct use values and/or non-use values;
- A wide range of values when using payment ladders (informed by the results of survey testing) to reduce anchoring and range bias;
- The use of entreaties such as cheap talk and oath scripts to reduce hypothetical bias.

Second, while momentary wellbeing indicators like happiness and purpose display intuitive associations with activities in the institutions, the life satisfaction approach does not yield meaningful estimates in the current study.

Third, the hybrid contingent/wellbeing valuation approach, which takes the hypothetical setting of CV and combines it with the underlying theory of WV, works well in eliciting plausible willingness to accept values. This method is useful where willingness to accept is a more appropriate measure of value than willingness to pay. For instance where there are sensitivities in asking visitors about their willingness to pay for entry or for a donation, our hybrid approach points to a sound new methodology that cultural institutions can use.

In this study, we work with two cultural institutions with very different missions and objectives to understand the applicability and limitations of economic valuation techniques in wide-ranging cultural contexts. In the contingent valuation we also employ different valuation scenarios and payment vehicles, as well as different population groups, to test its versatility. For these same reasons, however, this means that the valuation estimates for the two institutions strictly cannot be compared. Value transfer between institutions is only valid in cases where the institutions, the changes to be valued and the populations affected are similar, which is clearly not the case in the two institutions considered here. For the current purposes, therefore, our two case studies should be seen as separate, distinct and non-comparable examples of applicability of stated preference and wellbeing valuation methods in cultural institutions.

As a methodological contribution, we hope that our paper progresses policymakers' understanding of the different ways in which economic valuation techniques can inform the evaluation of the work of cultural institutions. We also intend the report to serve as practical guidance for institutions who are making the case to funders and policymakers for public investment.

#### **4.7. Interpreting the results further: implications of the results for areas outside of valuation**

There are a number of important messages to come out of this research in addition to the values estimated and the comparisons that are made possible between the two methods.

First, direct engagement with cultural institutions is associated with high levels of momentary wellbeing. In comparison to non-cultural activities, activities undertaken within the NHM and TL are associated with substantially large improvements in momentary self-reported wellbeing measures such as happiness and sense of purpose.

Second, these results for wellbeing are supported by evidence from methods that elicit preferences regarding different aspects of the work of the cultural institutions. The CV study results clearly show that people have an evident preference for use and non-use aspects of the cultural institutions.

Third, people who do not visit or directly benefit from cultural institutions value the existence of the institutions. This implies that cultural institutions have an important role in the quality of life of non-users and non-visitors as well as visitors in the UK.

#### **5. Future research**

Our study highlights important future areas of research. The detected life satisfaction impacts from engaging with cultural institutions in this study are different to those found consistently through analysis of cultural engagement in large-scale national datasets, such as the Taking Part Survey. This could however be due to a number of reasons, including the survey design. Future research could also develop longitudinal data to explore differences between the short-term and long-term wellbeing impacts of cultural engagement. This, however, still leaves open the issue of causality and so there is also a need to apply experimental methods like randomised controlled trials, to better capture unobserved idiosyncratic factors affecting the data, and to make stronger causal claims on the wellbeing value of cultural institutions like the Natural History Museum and Tate Liverpool.

## **1. Measuring Value in Cultural Institutions: Background and context of the project**

There are few more deeply contested relationships in cultural policy than that between economics and cultural value. Some argue that funders should use the techniques from welfare economics to value the non-market as well as market benefits of culture. Welfare economics approaches lie at the heart of Government decision-making. Capital expenditure in the cultural sector, as elsewhere, is reliant on cost-benefit analysis for making decisions amongst multiple spending options. As a consequence, economic valuation techniques have found support within the UK Department for Culture Media and Sport (DCMS) (O'Brien, 2010). Economic valuation approaches can also be used to make internal resource decisions within cultural institutions (the British Library study being a commonly cited example, see Pung et al., 2004).

In other cases it is advocated that funders take a more eclectic approach, using economic tools to value culture where it can be measured, but exploring alternative ways of capturing cultural value where it cannot (Throsby, 2001). Others question whether culture can be evaluated in welfare consequentialist terms at all (Tusa, 2000).

A striking feature of these debates in the UK is that they have rarely been informed by practical applications of economic valuation techniques (the most well known exceptions are contingent valuations by Jura Consultants, 2005; Maddison and Foster, 2003; Pung et al., 2004). Economic impact studies, by contrast, abound (Reeves, 2002). But impact studies invariably consider the secondary, market benefits of culture – the impacts on jobs, value added and exports etc; they do not provide an assessment of the value of the primary benefits of culture using the tools of public economics.

Arts funders have recently advocated a 'holistic' assessment of the benefits of culture, which goes beyond the economic and cultural to encompass the wellbeing, societal and educational value of culture (Arts Council England, 2014). But, absent of a theoretical framework, it is not clear how these different elements relate to one another, so when it comes to measurement there are risks of double counting.

Against this background, the AHRC's Cultural Value Project (CVP) signalled "a major interest in exploring the ways in which current valuation techniques recognised by the HM Treasury Green Book, the DCMS and other government departments are, or could with enhancement be made to be, fit for purpose in capturing, measuring and articulating cultural value".<sup>2</sup> Such techniques – by focusing on the primary benefits of culture – are also consistent with another of the CVP's aims, namely to prioritise the "more fundamental consequences of engagement with the arts and culture" giving particular "weight to how people experience various forms of art and culture in different contexts" (Crossick and Kaszynska, 2014).

The present study contributes to this agenda by applying two broad classes of technique that are recognised by the Green Book – preference-based and wellbeing valuation – to the work of two of the UK's premier cultural institutions, the Natural History Museum (NHM) and Tate Liverpool (TL) and drawing conclusions on the strengths and weaknesses of these methodologies when valuing culture.

The study is undertaken through primary data collection in a survey that includes a contingent valuation (CV) study and questions on subjective wellbeing (SWB). The closest antecedent study that directly compares preference-based and wellbeing valuation is Dolan and Metcalfe's (2008) valuation of an urban regeneration project, but since this early study the methodology for wellbeing valuation has been developed and refined.

A comparative analysis of these valuation methodologies is important for a number of reasons. First, the cultural sector has generally been hesitant towards traditional economic approaches to valuation, but arguably seems more open to the idea of wellbeing valuation.<sup>3</sup> Is this difference justified? Second, a detailed comparison of the two methods may hold out the promise of a more fit-for-purpose approach to economic valuation in the cultural sector. Third, the growth in interest in wellbeing valuation as an alternative valuation method more generally is particularly noticeable in the UK and in Europe. A rigorous investigation into its applicability in the cultural sector, and an understanding of how it differs from more traditional methods from cultural economics, is timely.

The research also contributes to the literature in a number of ways. First, it is the first comparison of stated preference and wellbeing valuation in the cultural sector and amongst the first ever conducted in any sector. Second, it provides the first

<sup>2</sup> [www.ahrc.ac.uk/Funding-Opportunities/Documents/CVP-Targeted-Call-Critical%20Reviews-Research-Dev-Awards.pdf](http://www.ahrc.ac.uk/Funding-Opportunities/Documents/CVP-Targeted-Call-Critical%20Reviews-Research-Dev-Awards.pdf)

<sup>3</sup> See, for example, [www.happymuseumproject.org](http://www.happymuseumproject.org).

monetary valuation estimates of access to and services provided by the NHM and TL through a comprehensive stated preference investigation, using state-of-the-art developments in the field. Third, it offers the first comprehensive subjective wellbeing investigation and valuation of NHM and TL visitation using a number of complementary wellbeing approaches.

With all this in mind, our primary research questions in this paper are as follows:

*Can preference-based and wellbeing valuation methods deliver convincing valuations in cultural institutions?*

*How consistent or conflicting are the valuations derived from applying these two methods?*

*Can techniques from the two valuation methods be combined to optimise their use by cultural institutions and funders?*

Section 2 of this paper provides an overview of key theoretical concepts for the research, including the advantages and disadvantages of preference-based and wellbeing valuations. Section 3 reviews the extant literature on these two approaches as applied to museums. The remainder of the paper is divided into two case studies. These sections have been written so that they can be read as self-standing reports.

In case study one, Section 4 begins by introducing the background on the cultural institution under study, in this case the Natural History Museum. Section 5 discusses survey design and application. Section 6 sets out the results and Section 7 offers an interpretation of them.

For case study two, Section 8 gives some background information on Tate Liverpool. As in case study one, Section 9 discusses survey design and application, section 10 sets out the results and section 11 interprets them.

Section 12 revisits the project's primary research questions, draws overarching conclusions from across the two case studies, and makes recommendations for valuation practice.

## **2. Economic valuation in the context of cultural institutions**

Understanding the economic value of a cultural institution is a complex issue. On the one hand, the activities it undertakes are wide-ranging and many of the services it provides are non-market in nature (most are free of charge). On the other, economic value has many dimensions and there are different ways of measuring it.

Cultural institutions produce goods and services that are different to those exchanged in regular markets. Their supply is fixed in the short run, and they are often publicly owned and freely accessible. Even when visitors are charged for entrance and/or for the use of services provided, the fees are usually nominal and not related to the true cost of providing services and maintaining the collections. This means that *non-market valuation methods* must be used to determine the value that people place on visiting these sites, using their services, and conserving them (Navrud and Ready, 2002). The importance of cultural assets in terms of their aesthetic, archaeological, educational, intellectual, collective identity, artistic and financial contribution has hence motivated an increasing application of valuation methods with the aim of deriving estimates of monetary values of cultural institutions from a societal point of view, which in turn could be used in project appraisal and decision making.

The economic value of something, as defined in the microeconomic theory that underlies HM Treasury's Green Book guidelines on cost-benefit analysis (CBA), is a measure of the change in human welfare, or *utility*, that results from it. We can define economic value in terms of primary and secondary benefits. **Primary benefits** are those that impact directly on an individual's utility, for example the educational benefits from visiting a museum or gallery. **Secondary benefits** relate to induced impacts felt by society more widely. These include, though are not restricted to, impacts on the economy and on the public purse. They could be in the form of increases in value added reflecting, say, the income earned by museum or gallery workers or from the profits made from museum or gallery-related tourism. These types of benefit are sometimes described as the 'economic impact' of a museum or gallery.

The focus of this study is on the primary benefits of cultural institutions, though we acknowledge that they are only one dimension – albeit a key one – of its overall value. That is, we value the impact of engagement with the cultural institution under study and the existence of the cultural institution on the public's utility.

Employing a Total Economic Value (TEV) framework (Pearce et al. 2006) – a typology of why individuals hold values for any given good, service or resource – we can identify a primary categorisation of use and non-use values associated with cultural institutions (Dasgupta and Serageldin, 1999, 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. Increasingly there may also be *distant use* benefits, generated by media, photographic or artistic representation of collections and assets as well as by digital engagement. *Indirect use* benefits could arise in the form of enhanced community image, and social interaction. A so-called *option value* can also be attached to potential future use of collections or of the services the cultural institution provides.

**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 renowned UK institution).

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 will be very 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. In our study we assess both use and non-use values associated with cultural institutions.

In this context, the study undertakes an explicit comparison between two methods of economic valuation in the context of: a) a large national cultural institution; and b) an important regional art gallery, using: (i) contingent valuation and (ii) wellbeing valuation. Before describing these methods in more detail it is useful to set out a formal framework for economic valuation.

## 2.1 The theory of economic valuation

As noted above, the value of a good or service relates to the impact that it has on human welfare (Freeman, 2003). We can express this in terms of compensating or equivalent measures as first devised by Hicks (1934).<sup>4</sup>

*Compensating surplus (CS) is the amount of money, to be paid or received, that will leave the agent in his initial welfare position following a change in the good.*

*Equivalent surplus (ES) is the amount of money, to be paid or received, that will leave the agent in his subsequent welfare position in absence of a change in the good.*

In essence, CS and ES refer to the change in income that holds welfare constant in light of the change in the provision of the good. This will be positive for goods that are welfare-increasing and negative for goods that are welfare-decreasing.

Hicks, in his pioneering work on the theory of value, did not initially propose a specific measure of welfare to be used in calculations of CS and ES. And so how these measures of value and welfare change would be assessed in reality was not clear until economists started to adopt a standard measure of welfare in empirical work. This came to be the preference satisfaction account of welfare.

### 2.1.1 Preference-based valuation

The preference satisfaction account of welfare states that “what would be best for someone is what would best fulfil his desires” (Parfit, 1984, p. 4). It rests on a set of rationality assumptions.<sup>5</sup> These assumptions allow us to map choices over a number of binary options on to a well-defined utility function which demonstrates that people behave *as if* they are maximising some utility function. Through these rationality assumptions preference satisfaction and welfare become synonymous with each other.

Traditional economic theory holds that utility is not observable, but measures of welfare change and monetary value can in theory be elicited from the expenditure function as follows.

<sup>4</sup> Definitions from Bockstael and McConnell (1980).

<sup>5</sup> These assumptions are (i) **completeness** (individuals are able to express a preference for any good or say they are indifferent between any pair of goods); (ii) **transitivity** (individuals who prefer (or are indifferent to) good x over good y, and who prefer (or are indifferent to) good y over good z, must also prefer (or be indifferent to) x over z); and (iii) **reflexivity** (individuals are indifferent between x and x). Note that for the purposes of valuation we need to add two further assumptions (Freeman 2003): (i) **non-satiation** (that preferences are never fully satiated such that the individual always places a positive value on more consumption); and (ii) **substitutability** (if the quantity (or quality) of one good decreases it is possible to increase the quantity (or quality) of another good sufficiently to make the individual indifferent between the two states of the world).

The individual's *primary* problem is to find the solution to the (direct) utility function:

$$x \max U(X, Q) \quad (1)$$

which is to find the optimal level of consumption of market goods ( $X$ ) to maximise utility ( $U$ ) given a budget constraint and a certain level of non-market goods provision ( $Q = Q^0$ ).

Under the duality principle the indirect utility function derived from the direct utility function can be represented through the expenditure function:

$$E = E(p_x, Q, U^*) \quad (2)$$

which shows how expenditure changes as a function of the prices of market goods and provision of the non-market good, such that the individual continues to maximise utility,  $U^*$ .

The expenditure function circumvents the problem of unobservable utility as it allows us to calculate the amount of money we need to *give* or *take away* from the individual to make her equally well-off after the policy or provision of the non-market good. This is the difference ( $E_{post} - E_{pre}$ ), where  $E_{post}$  = the level of expenditure *after* the change in the non-market good (here assume the quality or quantity of the non-market good has *decreased*) and  $E_{pre}$  = the level of expenditure *before* the change in the non-market good. So in this example the difference ( $E_{post} - E_{pre}$ ) is the amount of money required to compensate the individual for the drop in the non-market good.

Under different formulations of the expenditure function, the difference ( $E_{post} - E_{pre}$ ) is the CS or ES of the change in  $Q$ . CS and ES are often rephrased in terms of **willingness to pay (WTP)** or **willingness to accept (WTA)** and Table 2.1 describes the relationship between these concepts of value.

**Table 2.1. The relationship between CS, ES, WTP and WTA**

	Compensating Surplus (CS)	Equivalent Surplus (ES)
<b>Welfare gain</b>	WTP for the positive change	WTA to forego the positive change
<b>Welfare loss</b>	WTA the negative change	WTP to avoid the negative change

Generally speaking, where proxy markets exist the favoured approach is to estimate WTP or WTA from people's market behaviour using **revealed preference** valuation methods. Revealed preference methods uncover estimates of the value of non-market goods by using evidence of how people behave when presented with real choices. The basic idea is that non-market goods affect the price of market goods in other well-functioning markets and price differentials in these markets can provide estimates of WTP and WTA. Revealed preference methods relevant to the cultural sector include:

- (i) the travel-cost method, which has been used to elicit mainly recreation benefits estimated from the actual travel expenses incurred by individuals when visiting a specific cultural site; and (ii) the hedonic pricing method, which elicits the value of a non-market good as part of the attributes' bundle of a market good e.g. house property prices which may vary due to proximity to cultural institutions (Pearce et al. 2006).

Often, however, proxy markets do not exist for the non-market good in question, and in these situations economists have traditionally turned to direct elicitation of WTP and WTA through survey-based **stated preference** valuation methods. Our focus in this study is on contingent valuation. The **contingent valuation method (CVM)** is a stated preference survey-based methodology that seeks to elicit monetary values for non-market goods by directly asking individuals about their willingness to pay or willingness to accept a particular change (Bateman et al. 2002). Respondents are presented with a hypothetical market that describes in detail the proposed change under valuation, using baseline conditions as a reference point. The hypothetical scenario should be understandable, plausible, and meaningful to respondents so that they can give valid and reliable values despite possible lack of experience with one or more aspects of the scenario (Mitchell and Carson, 1989). Respondents are then asked how much they would be prepared to pay (accept) for the change described.

The valuation questions can be presented in a number of different ways, including open-ended, bidding game, payment card, and dichotomous choice elicitation formats.

**Open-ended questions** represent the simplest direct elicitation technique where respondents simply state the amount they would be willing to pay. However, a number of problems have been associated with this technique, including large non-response rates, protest bids, and outliers that skew the data (Bateman et al., 2002; Mitchell and Carson, 1989). At the opposite

end of the spectrum, **single-bounded dichotomous choice questions** allow the researcher to randomly present a monetary figure to the respondent (Arrow and Solow, 1993). This simplifies the cognitive task and is thought to provide incentives for truthful revelation of WTP (if their WTP is higher, say, than the amount offered, it is in their interest to accept this value). However, this approach may also encourage a larger number of protest bids than open-ended mechanisms. It also provides data only in a range above or below which we can be certain the respondent is willing to pay. **Double-bounded dichotomous choice** improves on this method by asking a follow-up question: if the respondent indicates that they would pay the initial amount (e.g. £2), would they pay the next amount up (i.e. £3)? Alternatively, if the respondent indicated that they would not pay the initial amount of £2, would they be willing to pay the next amount down (i.e. £1)? Although this improves on the single-bounded dichotomous choice approach, both mechanisms introduce so-called anchoring bias, by suggesting to the respondent the appropriate range of values to be elicited (Arrow and Solow, 1993). **Payment card** approaches present respondents with a range of monetary amounts from which they are asked to pick their willingness to pay. This eliminates anchoring (and starting point) bias and provides a visual aid to the cognitive process of valuing the good (Bateman et al., 2002; Maddison and Foster, 2003; Maddison and Mourato, 2001). However, this approach is in turn vulnerable to biases by the range presented to the respondent. The inclusion of an open-ended 'other amount' option contributes to reducing this effect.

Contingent valuation questionnaires also normally contain a wealth of additional questions to gain information on a respondent's socio-demographic characteristics, their attitudes towards the good, and the reasons behind their stated valuations. The responses to these questions are typically used to model the determinants of WTP (or WTA).

Finally, although not the technique used in this report, it is worth noting that there is another well-known stated preference technique, generally known as **choice modelling (CM)**. Choice modelling is applied by economists in situations that are multi-dimensional in nature and where it is important to find the value attached to the various dimensions of interest. Respondents are presented with a series of policy scenarios, grouped into 'choice sets' that describe the alternatives on offer. The scenarios are described in terms of a number of attributes that vary at different levels, between scenarios. One attribute typically represents a monetary variable, which enables the derivation of implicit prices (i.e. WTP). Respondents are then asked to identify their most preferred scenario, amongst the scenarios contained in a choice set. One of the scenarios typically describes a current or future status quo option (the reference point), and remains constant between the choice sets. Willingness to pay (or to accept) is therefore inferred indirectly by analysing the choices and trade-offs made between the various attributes. For a review of the advantages and disadvantages of CM vis-à-vis CV see Hanley, Mourato and Wright, 2001).

### 2.1.2 Wellbeing valuation

The premise of the **wellbeing valuation (WV)** approach is to estimate measures of welfare change (CS and ES) from data on people's experiences as measured by their SWB. SWB is based on people's self-reports about how their lives are going. This is normally in the form of a single-item question such as "how happy are you now?" (an example of a momentary wellbeing measure) or "overall how satisfied are you with your life?", where responses are measured on numeric scales such as 0-10. This form of welfare measurement gets around the issues of adherence to a strict set of rationality assumptions as in the preference measures that underlie the CV approach, but raises the question of whether questions on SWB – be they momentary or related to life satisfaction – are able to provide accurate reflections of one's quality of life, an issue that is discussed in more detail below.

In WV, SWB is assumed to provide a direct measure of an individual's welfare and hence a key assumption is that welfare is now observable as measured by SWB. This allows us to estimate compensating and equivalent measures of value using the direct utility function (where utility is now some measure of SWB) without recourse to the duality principle and the expenditure function:

$$SWB(Q, M) \quad (3)$$

Where  $SWB(\cdot)$  is a direct SWB function in which  $Q$  = the non-market good and  $M$  = income. In its most basic format, the value (here CS) of the non-market good in the WV method is estimated as:

$$SWB(Q^0, M^0) = SWB(Q^1, M^1 - CS) \quad (4)$$

Where superscripts 0 and 1 respectively signify conditions before and after provision of the good,  $Q$ , which in this exposition is assumed to have a positive impact on utility.

Measures of welfare change can be estimated from the marginal rates of substitution (MRS) between the non-market good and money in the SWB function and this is estimated empirically, by for example using the coefficients from a regression model. The statistical methodology underpinning the WV approach is set out in full in Sections 5.2.2, 9.2.2 and Annex C.

The WV method has been used to value goods, 'bads' and outcomes, as wide-ranging as the environment and environmental amenities (e.g., Carroll et al., 2009; Ferreira and Moro, 2010; Rehdanz and Maddison, 2008; Welsch, 2002, 2007; Welsch and Kühling, 2009); health (e.g., Ferrer-i-Carbonell and Van Praag, 2002; Groot and van den Brink, 2006); crime (e.g., Cohen, 2008; Moore, 2006); public sector corruption (e.g., Welsch, 2008); civil conflicts (e.g., Welsch, 2008); care-giving (van den Berg and Ferrer-i-Carbonell, 2007); terrorist attacks (e.g., Frey and Luechinger, 2004); housing quality (Fujiwara, 2013a); social relationships (e.g., Powdthavee, 2008); employment (e.g., Clark and Oswald, 2002); macroeconomic events (e.g., Blanchflower and Oswald, 2004); value of life (Deaton, 2008; Oswald and Powdthavee, 2008); commuting (Stutzer and Frey, 2004); adult learning courses and qualifications (Dolan and Fujiwara, 2012); income inequality (Beja, 2011); and cultural activities and events (Fujiwara, 2013b; Fujiwara et al., 2014).

CV and WV provide two distinct ways of estimating economic value based on two substantively different approaches to measuring human welfare: preferences and SWB respectively. As such, we would not expect them to give the same results – for example, it may be that people desire things that do not impact positively on their SWB. Our study assesses the extent to which the valuation results from these methods are comparable and whether there are any clear advantages of one approach over another in the specific context of valuing the work of a cultural institution.

## 2.2 Comparing contingent valuation and wellbeing valuation

There are a number of comparative strengths and weaknesses of the CV and WV approaches which must be borne in mind.

### 2.2.1 Advantages and limitations of contingent valuation

The key methodological advantage of CV over other non-market valuation approaches is that it can measure values and benefits that would not be revealed under market conditions, such as *non-use* values. Importantly, it also offers the opportunity to measure benefits associated with changes that have not yet happened, that is, *future* changes (such as those related to a cultural institutions' conservation work, or investments which mitigate irreversible risks).

Because CV asks directly about hypothetical scenarios, it is able to deal with marginal changes in a good or service such that a wide range of specific events and activities can be valued. Relatedly, it is possible to better pinpoint the item or service being valued by being more specific in the survey (this is discussed in more detail later).

A further significant advantage of CV over WV is that it is not reliant on complex statistical analysis. Indeed, WTP values can be estimated with minimal statistical analysis (usually just by deriving average values if samples are representative) and statistical analysis only becomes necessary when analysing the drivers of WTP values.

A practical advantage that CV has over WV is that it has a significantly longer history of research in economics. In particular, the rise in popularity of stated preference methods in the early 1980s was accompanied by a very active debate and critical assessment of the merits and limitations of the techniques and their underlying conceptual framework. According to Kerry Smith (2000), "*Contingent valuation has prompted the most serious investigation of individual preferences ever undertaken in economics*". As a consequence, a lot more is currently known about the problems of this particular method, ways of testing for them and techniques for addressing them.

Against these advantages, there are a number of well-known potential biases in CV that can be problematic if not adequately addressed in the survey instrument and analysis (Bateman et al, 2002, Pearce et al., 2006, Fujiwara and Dolan, forthcoming). These include: **insensitivity to scope**, where WTP is insensitive to the scope of the proposed change; **protest values**, where respondents have a principled objection to providing a monetary value; and **strategic bias**, where respondents seek to 'game' the study by providing values that they think will influence an actual resource allocation recommendation. Finally, differences between WTP and WTA can also occur, which violates the underlying theory of economic preference satisfaction (Hausman, 2012).

The criticism of CV that has perhaps received greatest attention is **hypothetical bias** (Arrow and Solow, 1993; Champ and Bishop, 2001; Hausman, 2012), where individuals' stated WTP may be significantly larger than actual WTP due to the hypothetical nature of the survey. This arises mostly when a *voluntary payment mechanism* is used. A range of tools have been developed to try and correct for this (Sugden, 2005). Champ and Bishop (2001) tested the use of so-called *certainty questions* to control for the bias ("how certain are you that you would really pay the amount indicated if asked"). Specifically, they recoded the positive responses of less certain respondents as negative responses in order to derive a lower bound estimate of the compensating surplus. The removal of 48% of those who responded affirmatively but uncertainly to the contingent donation question appears to have eradicated significant hypothetical bias. In other words, once uncertain respondents were removed, the contingent donors could not be statistically distinguished from actual donors across a wide range of characteristics.

A common way of addressing hypothetical bias is through the use of so-called *ex ante entreaties* in the survey instrument. Famously, Cummings and Taylor (1999) developed a *cheap talk* entreaty for reducing hypothetical bias associated with provision of a local public good. Cheap talk consists of a script describing the bias problem and a plea to respondents that they do not overstate their true willingness to pay. Subsequent work on CV and discrete choice models has found mixed results from the use of cheap talk, however (e.g., Atkinson et al, 2012; Aadland and Caplan, 2006; Carlsson et al., 2005; Carlsson and Martinsson, 2006; List and Lucking-Reiley, 2000; Lusk, 2003). For example, Murphy et al. (2003) tested the robustness of the cheap talk procedure using a modified *voluntary contribution mechanism* (VCM) (see Rondeau et al., 1999) incorporating a one-shot provision point with a money-back guarantee if the provision point is not met. Respondents were asked to rate on a scale of 1 to 10 the extent to which the following statement influenced their decision about how much to pay: "Since others will pay, I do not need to contribute as much." Subjects were also asked to indicate on a scale of 1 to 10 the degree to which they agreed with the statement that "most other people will contribute less than what (the good) is really worth to them because they will be able to enjoy the benefits regardless of how much they contribute." Respondents were also asked how much their decision about how much to pay depended on what they thought others would do, and if they would change their payment decision if they knew exactly what others had paid. Cheap talk reduced the percentage of subjects giving yes responses relative to subjects in the hypothetical situation without cheap talk. But, unlike Cummings and Taylor (1999), cheap talk did not eliminate the difference between real and hypothetical responses: hypothetical mean WTP exceeded actual payment by a factor of about 7.6:1. Although the cheap talk script reduced this difference to 4.8:1, the reduction was not statistically significant.

Morrison and Brown (2009) tested three techniques for reducing hypothetical bias: certainty scales and cheap talk – as discussed above – and *dissonance minimization*, which adds response categories that permit respondents to express support for a program without having to vote in favour of increased expenditure. They found that certainty scales – when properly calibrated – and dissonance minimization were the most effective in reducing the bias.

While cheap talk has been one of the most popular strategies to reduce hypothetical bias, more recently another entreaty has been developed: the *oath script*. The oath script was proposed as a potential addition or alternative to the cheap talk script. Typically, the oath script asks respondents to agree to promise (or swear a solemn oath on their honour) that they will respond to questions or state values honestly. Within economics, however, the oath script has seen very few applications outside a lab setting. In fact, there have been only a handful of field studies altogether using the oath script in non-market valuation (Ehmke, Lusk, & List, 2008; Carlsson et al., 2013).

A particular form of hypothetical bias, though one which could in principle go either way, reflects **focusing illusions**, whereby at the time of preference elicitation, people focus only on the salient aspects of the outcome (or condition) and this may not reflect in any way how people would actually experience these conditions or states in real life (Hausman, 2012; Sugden, 2005). People may not in any case be able to predict how much they will in fact like (or dislike) something.

Problems also arise with CV when people have insufficient information to make informed choices, and a growing literature shows that preferences can in this case be influenced by contextual factors. For example, a number of studies have shown that it is possible to reverse people's preferences and that WTP values can be primed during a survey. Other research shows that WTP values can be influenced by random numbers (known as anchors or primes) present in the individual's environment during the survey (Ariely, 2000) and by the pleasantness and smell of a room (Poundstone, 2010). The implication for valuation theory is that what people want or prefer may not always align well with what is best for them.

Last, but not least, some have raised ethical concerns about the appropriateness of using WTP at all to value services like health and culture. This may, for example, be because value is related to ability to pay (and the prevailing income distribution may be seen as inequitable); or because using money to value health and culture may send an undesirable signal (namely that health and cultural services are just like any other commodity bought and sold in the market place) (Fujiwara and Dolan, 2014).

### 2.2.2 Advantages and limitations of wellbeing valuation

WV, in contrast, avoids problems related to protest values, strategic bias, hypothetical bias and contextual influences such as priming effects, because WTP is not elicited. Importantly, WV also eliminates focusing illusion issues, since econometric analysis is conducted to assess the importance of a factor in someone's life without having to ask them to attribute a value to that factor. The WV approach is based on real experiences and not – as in stated preference approaches – on how people *imagine* they might be affected by a change. In wellbeing analysis, it is possible to look at people's actual experiences of engaging with a cultural institution when they are living life as they normally do, and attach values to these conditions. In this sense WV is more akin to a revealed preference technique. This has an additional advantage: people are not asked to place a value on the non-market good directly, so it avoids ethical difficulties which arise when respondents are asked to assess the value of a cultural good or service in monetary terms.

Such factors no doubt help to explain why Dolan et al. (2011) found that from among a range of quality of life measures, including wellbeing and preferences, people generally favoured evaluative measures of wellbeing (such as life satisfaction) as the basis for government decision making and policy evaluation in the UK. Although this study was conducted with a small sample, the results suggest that SWB has a role to play in public policy decisions (see also O'Donnell et al, 2014).

However, there are concerns that the SWB scores which underpin WV are insufficiently sensitive to small differences in actual wellbeing due to the bounded nature of SWB measures (Johns and Ormerod, 2007; Tella and MacCulloch, 2006). This may be of particular concern with impacts from museum or gallery visits or other discrete activities where the wellbeing impact might be expected to be limited. Furthermore, after some point a wellbeing impact – no matter how large – may not show up in people's self-reports, as these are constrained by the upper and lower bounds of the SWB scores.

A second limitation with WV is that it cannot be used to measure non-use values in any obvious way, since the individual needs to have experienced ('used') the good or service for it to show up in their wellbeing responses. In principle, they could be assessed by providing individuals with information about a change in circumstance about the good or service which affects their wellbeing (despite no planned use). For example, Metcalfe et al. (2011) found that (information about) the 9/11 terrorist attacks had a large effect on measured wellbeing in the UK. However, estimation of non-use value is not something that has been explored in the WV literature.

Subjective wellbeing has its own measurement challenges. Life satisfaction, for example, can be seen as being made up of a balance of affect (positive and negative emotions and feelings) together with a cognitive assessment of how well one's life measures up to aspirations and goals (Diener, 1984; Kahneman and Krueger, 2006). In other words, a life satisfaction response incorporates a characterisation of how one feels now and a retrospective judgement of one's life (Kahneman and Krueger, 2006). This retrospective element is problematic, because people have **imperfect recollection** of past experiences (e.g. if they follow Kahneman's peak-end rule, whereby individuals judge an experience by its most intense point and its end). Furthermore, their present feelings can be influenced by **contextual factors** present at the time of survey, and biases can also arise in the stage of verbally reporting life satisfaction scores (Bertrand and Mullainathan, 2001; Kahneman and Krueger, 2006; Loewenstein and Ubel, 2008; Schwarz and Strack, 1999). For example, life satisfaction scores have been found to be influenced by diverse factors such as the weather, finding a dime on a copying machine and whether a football team one is watching wins (Schwarz and Strack, 1999). While such factors might plausibly be expected to influence current mood, they should not have notable effects on global or overall measures of SWB, such as life satisfaction. Fundamentally, single-item measures such as life satisfaction and happiness may not be broad enough to tap into or reflect all that is important to our lives (Loewenstein and Ubel, 2008).

A particular problem in the WV approach is that it heavily relies on statistical analysis, which when employed on non-experimental data makes it difficult to assess cause and effect. The essential step in WV involves estimating the causal impact on subjective wellbeing of the non-market good or service in question and of income (which, as we will see, allows wellbeing impacts to be translated into monetary terms). This involves a number of problems, including **selection bias, reverse causality** and **measurement error**. A common outcome is that values estimated using WV tend to be too high because the income coefficient (the impact of income on life satisfaction) is downwards biased (Dolan et al., 2008). For example, in some studies the value of employment has been estimated to be an implausibly high £23,000 per month to the individual in addition to their wage income (Clark and Oswald, 2002), and the costs associated with drug and alcohol problems have come out at as high as £9 million per year (Powdthavee and van den Berg, 2011) (This issue is discussed in more detail in Annex C, together with some solutions that have been proposed in the literature). It can also be difficult to pinpoint accurately the results in the statistical analysis. For example, a positive coefficient on a dummy variable for living near a park could be used to value access to the park using WV, but it could have many other interpretations too.

In sum, the literature highlights a number of comparative advantages and disadvantages associated with CV and WV approaches which precludes any prior judgement on which works 'best' when applied in the cultural sector. Hence the focus in this paper on an empirical examination of the issues in the context of two practical case studies, of the Natural History Museum and Tate Liverpool.

Before we present our case studies we present a brief review of the recent literature on economic valuation in museum and galleries.

### **3. Literature review of valuation studies in museum and galleries**

Empirical research eliciting economic values or benefits associated with, access, preservation or restoration of cultural assets dates back to the 1980s when the first valuation studies in the field were conducted, focusing on the arts, theatre, historical sites, museums, galleries, libraries and broadcasting (Noonan 2003). Since then, many studies in the cultural sector have been conducted worldwide investigating a variety of benefits, both tangible and intangible. The majority have used preference-based techniques, mostly stated preference methods and in particular CV, and have aimed to elicit visitor and/or non-visitor – that is, willingness to pay – values. But in recent years, studies using WV have also begun to appear. Detailed reviews of stated preference heritage valuation studies include Noonan (2003), Pearce *et al.* (2002), and Provins *et al.* (2010). However, none of these reviews look at wellbeing valuation studies.

We distinguish between studies that assess use value, in terms of the direct use arising through experiences of the cultural institution, and non-use values, in terms of the existence of the institution, the bequeathing of an institution for future generations, or the option value of knowing one will have the possibility of engaging with the institution in the future even if one has not used that option yet (Bateman *et al.*, 2002; Mitchell and Carson, 1989; Throsby, 1999). In order to capture the total costs and benefits of a cultural institution, the full range of non-market effects have to be taken into account.

In this paper, we review valuation studies with a focus on museums and galleries, covering those included in previous reviews as well as more recent research. Our review spans the use of the full range of valuation methods, not just stated preference methods. For ease of interpretation in all cases we convert foreign currency to GBP at the time the study was published using the relevant Consumer Price Index (CPI), and then convert this amount to present-day prices using annual Retail Price Index inflation (figures based on the RPI as of May 2014. Source: Office for National Statistics).

#### **3.1 Stated preference studies**

Our review of the museums and galleries valuation literature shows that UK applications are in fact the most frequent, including two applications to the British Museum (a CV and a CM). Spanish, Italian and Scandinavian applications follow in number. There is a good deal of variety in terms of the types of value assessed, with studies on preservation and conservation being the most popular (six applications), followed by five studies considering increases or changes to the provision of an institution's services. The majority of studies value both use and non-use values (five applications), four studies assess use value only and three measure non-use value only. The majority of studies focus on donations or voluntary contributions to a fund as the CV's payment mechanism, with only very few using entrance fees or taxes. The most common elicitation mechanism is seemingly the dichotomous choice question, used in six studies. Prior to 2002, single-bounded dichotomous choice is found in three studies. After 2002, the most common elicitation mechanism is the double-bounded dichotomous choice with open-ended questions. Payment ladders are used in fewer applications.

Santagata and Signorello (2000) conducted a general population survey of 468 respondents to measure the non-use value of a network of cultural and historic monuments making up the Napoli Musei Aperti in Italy. The hypothetical scenario was that public funds (local taxes) would be withdrawn and substituted by a non-profit operator. Two elicitation mechanisms – single-bounded dichotomous-choice bid and open-ended questions – were tested. The authors reported mean WTP values of 17,000 lire (£8.37 present day GBP) and 30,000 lire (£14.76 present day GBP) from the open-ended and dichotomous-choice questions, respectively. They argued that the disparity between elicitation methods was unlikely to be caused by strategic bias, since the incentive for understatement of the true WTP should have been modest in the case of donations. Instead, they attributed the difference to the cognitive difficulty and preference uncertainty of open-ended elicitation, making lower values more likely, and the effect of yea-saying responses to the dichotomous choice question.

Maddison & Foster (2003) valued congestion costs in the British Museum with a photomontage survey associating crowded conditions with free admission (the status quo) and less-crowded photos with an admission charge. Dichotomous choices of £3, £6, £12 and £20 were randomly presented to a sample of 400 museum visitors. The estimated congestion cost imposed by each additional visitor turned out to be £0.05 (£0.07 present day GBP). Multiplied by the daily average number of visitors this gave an implied aggregate congestion cost for each visitor of £8.05 (£11.53 present day GBP). The authors note that, unlike WTP, there is "no necessary link between the marginal congestion cost and the optimum charge", since imposing this congestion charge would lead visitor numbers to fall, in turn reducing the congestion externality (Maddison and Foster, 2003, p. 186).

Sanz *et al.* (2003) elicited willingness to pay for the preservation and maintenance of the National Museum of Sculpture in Valladolid, Spain. Using the payment mechanism of an annual donation to a preservation fund, they captured use value, in the form of 1,108 on site surveys, and non-use values through 1,014 telephone surveys with local residents. Use of a double-

bounded dichotomous choice elicitation method gave a mean direct use value WTP of between 25EU and 30EU (£23-28 present day GBP) under a conservative scenario, and between 33EU and 40EU (£30.75-37.25 present day GBP) under a more optimistic scenario. Non-use value was estimated to be approximately 27EU and 36EU for each of these scenarios (£25-33.50 present day GBP). The authors reported no great difference between the valuation of the direct users and non-users, and in some cases non-user values were even 2EU higher than the estimate of the value to direct users.

The British Library (BL) study (Pung et al., 2004) is commonly held up as a best practice example in CV analysis. It presented a hypothetical scenario that government funding for the BL would cease to be provided. The survey provided a payment mechanism of a donation or subscription to help support the continuation of the reading rooms and remote services. The researchers used a two-stage bidding game elicitation method on 2,539 reading room users and a general public non-user sample of 2,030.

The BL study also collected users' willingness to accept data for a scenario where the BL ceased issuing readers' passes but allowed existing readers to sell their pass, asking the minimum amount respondents would be WTA as a monthly payment in compensation. They also gathered data on the incremental cost of alternatives for users, such as travel, accommodation and cost of access to materials, and a non-use question using a payment mechanism of raised taxes for the maintenance of reading rooms, remote document supply and bibliographic services.

Pung et al. (2004) estimated a mean direct use WTP of £116 (£162 present day) and a higher WTA value of £273 (£381 present day) in monthly compensation for reading room users. The general public non-use WTP was much lower, at £6.30 (£8.80 present day). Direct value amounted to £59 million (£82 million present day) and indirect value amounted to £304 million (£424 million present day). The authors therefore estimated the Total Economic Value of use and non-use of the British Library to be £363 million per year (£506 million present day).

A similar CV approach was applied to Bolton's museum, library and archive services, consisting of three museums, fifteen local libraries and a central archive (Jura Consultants, 2005). Using the scenario that funding from the local council would cease, the researchers presented the payment mechanism of a donation to support the continuation of the museums, libraries and archives again through a two-stage bidding game elicitation method. The study captured use and non-use value through a sample of 325 face-to-face and telephone respondents divided into the discrete categories of museum, library and archive users and non-users. They also asked the WTA monthly compensation to give up the museum, library, archive pass, and gathered data on travel time and cost of alternatives. Among users, the study found a mean annual WTP for museums of £33.24 (£44.94 present day), libraries of £39.96 (£54.03 present day), and archives of £21.96 (£29.69 present day). The total use value for Bolton's museums, libraries and archives was therefore estimated at £95.16 (£128.66 present day). For non-users, mean WTP for museums was calculated as £13.68 (£18.50 present day), libraries as £12.00 (£16.22 present day), and archives as £8.16 (£11.03 present day), giving a total non-use value of £33.84 (£45.75 present day). The authors aggregated WTP to £10.4 million (£14.06 million present day), divided between direct value to museum users of £2.8 million (£3.79 million present day), library users of £4.4 million (£5.95 million present day), and archive users of £0.2 million (£0.27 million present day), and indirect value of £3 million (£4.06 million present day).

Stevenson (2013) performed a contingent valuation study of the National Galleries of Scotland to estimate an overall consumer surplus provided by the National Galleries of £3,567,703. The authors note that although this surplus is primarily predicated on the benefits it provides to users as individual consumers, in arriving at their valuation, participants also factored in the societal and economic benefits they value as citizens.

There have been a number of CV studies valuing museums in other countries. Tomho (2004) measured the non-use value to 800 local residents of keeping the Museum of Central Finland in existence through annual taxes, using an open-ended elicitation mechanism. It estimated a mean WTP of FIM103 (£16.13 present day GBP) and median WTP of FIM50 (£7.83 present day GBP).

Lampi and Orth (2009) combined CV methods with an analysis of the socio-demographic effects of the reintroduction of entrance fees in Sweden in 2007. They used two surveys to collect information about the visitors' WTP and socio-economic backgrounds, one before the fee came into effect (from a sample of 565 visitors) and one after (from a sample of 315 visitors). The elicitation mechanism was an open-ended question, but the authors tested the validity of their results by investigating whether predicted changes in socio-demographic composition of visitors based on maximum WTP differed from the actual changes observed after the introduction of the entrance fee. The mean WTP during the free entry period (2006) was 56.6 SEK (£5.44 present day GBP), and the median was 50 SEK (£4.82 present day GBP). The results of the CV survey suggested that setting the fee at 40 SEK (£3.76 present day GBP) would lead to a significant decrease in the shares of visitors who were immigrants. To test the validity of the CV, the authors compared the characteristics of those who had visited the museum after

the introduction of the entrance fee and who therefore paid 40 SEK with those who visited the museum while entrance was still free and had a maximum WTP of at least 40 SEK. They found, in contrast to the predictions of the CV, the proportion of first-generation immigrants was significantly higher after introduction of the fee.

Bedate et al. (2009) measured willingness to pay for a cultural good, dependent on the level of certainty expressed by individuals in their responses. Using double-bounded dichotomous choice questions, the authors found WTP was lower when certainty was higher. At a certainty level of  $\geq 7$  (on a scale of 0-10), mean WTP was around 18EU (£23.19 present day GBP) for museum visitors. In contrast, when all moderately certain responses were coded as uncertain (an asymmetric uncertainty model) mean WTP was much lower, estimated at 14EU (18.04 present day GBP) for museum visitors.

Snowball (2005) calculated the non-use value for local residents for two arts festivals in South Africa, presenting the hypothetical scenario of a 25-50% reduction in the size of the festival. The total sample of 279 phone interviews were randomly assigned single-bound dichotomous choice followed by an open-end question. Upper bound mean WTP in Grahamstown for high income individuals was \$2.28 (£1.60 present day GBP) and for low income individuals \$1 (£0.70 present day GBP). Upper bound mean WTP in Oudtshoorn for high income individuals was \$2.68 (£1.88 present day GBP) and for low income individuals \$1.59 (£1.11 present day GBP). The overall WTP (upper-bound) for Grahamstown was \$462,000 (£323,317 present day GBP) and for Oudtshoorn \$446,000 (£312,120 present day GBP) per year.

Báez-Montenegro et al.'s (2012) study of the non-use value of cultural heritage in Valdivia, Chile also revealed a slight drop in WTP as the degree of certainty of actually making a payment, in accordance with the hypothetical valuations stated, was higher. At maximum certainty (53.18% of interviewees), estimated WTP (\$11,617; £7,500 present day GBP) was 29.3% lower than for the sample as a whole.

Burton, Louviere, and Young (2009) used choice modelling to value two Australian museums, using an online survey of 82 respondents for the national museum and 89 for the state museum. Eight discrete choice scenarios were developed, framed in terms of logistics (travel), visit duration, and cost. Respondents were asked to indicate whether Option A or Option B would result in more frequent visits, less frequent visits, or no change to visiting patterns. For the state museum, multinomial logit (MNL) choice models showed that significant factors were: summer opening hours; family evenings; chat with curator; after school programs; day ticket on monorail; combined entry to IMAX cinema; joint museum pass; guided walking tour; and re-entry on same pass. For the national museum, significant attributes included: extended summer opening hours; joint ticket with the Imax/Aquarium; harbour cruise; ferry tickets; joint membership and fee options; and re-entry in the same month.

Jaffry and Apostolakis (2011) applied discrete choice methods on 500 visitors to the British Museum. The choice modelling exercise measured willingness to contribute (WTC) for three described states of improvements or deteriorations from the fourth scenario (status quo). Attributes included: opening hours, temporary exhibitions, information and communication technology (ICT), staffing, facilities management, outreach and engagement, and voluntary contributions, reflecting the range of fees charged at museums worldwide and included as a proxy for 'price'. Survey respondents reported that they would be willing to contribute up to £2.60 for five temporary exhibitions per year. They were in general positive towards ICTs at the museum, reporting a £1.64 mean WTC for the provision of information after the end of their visit. For facilities management, respondents in general were prepared to contribute up to £1.50 for a bigger information desk, leading the authors to conclude that respondents' strong preferences towards more ICT provisions indicate that congestion levels may be an inhibiting factor for museum attendees.

Mazzanti (2003) applied discrete choice modelling to value the Galleria Borghese in Rome. 185 gallery visitors were presented with different scenarios in terms of access time, ancillary services and entrance fee versus the status quo. WTP was elicited for visiting the museum as it is (status quo) and for a special conservation fund for the museum. The payment mechanism was an admission fee/special conservation fund with a payment ladder elicitation method. Mean and median WTP for the visit were 7,460 and 6,000 lira respectively (£3.66-£2.96 present day GBP). Mean and median WTP for the conservation fund were 16,990 and 16,000 lira (£8.35-£5.14 present day GBP). The admission fee was 14,000 lira (£6.88 at present day GBP). Mean WTP for temporary exhibitions and multimedia services ranged from 4,300-5,900 lira (£2.12-£2.91 present day GBP). Mean WTP ranged from 6,600-7,470 lira (£3.25-£3.68 present day GBP) for special conservation activity. There was no WTP for increased access times.

### 3.2 Revealed preference studies

We consider two methods for inferring the value of museums and galleries (and other cultural institutions) from the preferences revealed by consumers.

The first approach involves estimating **how visitor numbers vary with the attributes of an institution or change following a discrete intervention**, such as an improvement in the customer experience or built environment (e.g., Alberini and Longo, 2006). This approach applies choice models to take into account individual differences in willingness to travel caused by the different contexts of visits (McFadden, 1977, 1974).

The second method estimates **average travel time**, measured in questionnaires by asking respondents for their postcode of residence, region of origin, or nationality (Bedate et al., 2004; Boter et al., 2005; Brida et al., 2012; Fonseca and Rebelo, 2010). For example, Zonal Travel Cost (ZTC) methods (Hotelling, 1949) are derived from the number of trips originating from a zone divided by the population of that zone. By calculating the average cost of the trip and the percentage of visits for each zone, a demand curve may be constructed, which represents the variation in the number of visits caused by changes in the cost of travel. Once the final demand curve is constructed, the area under the curve gives the consumer surplus. Revealed preference models must also control for individuals' budget and time constraints. They also need to overcome the difficulties in assigning costs to multiple sites visited on the same trip (Clawson and Knetsch, 2013). More recent papers have followed an alternative individual travel cost (ITC), which uses trips per year by individual users of a site as the dependent variable (e.g., Brida et al., 2012; King and Mazzotta, 2011). Finally, **hedonic demand models** – not reviewed here – reveal preferences towards a non-market good by decomposing the item into its constituent characteristics, obtaining estimates of the contributory value of each characteristic (Malpezzi, 2003; Rosen, 1974).

The first class of studies attempts to estimate the causal relations between museum attributes and visitor behaviour (that is, the number of repeat visits, visit duration, time spent for rests etc.), satisfaction, and mood. The definition of satisfaction in these studies includes cognitive elements, fulfilment of visitor expectations (Churchill & Surprenant, 1982; Oliver, 1980; Oliver & Swan, 1989; Prakash, 1984), and affective elements, created when cultural exhibitions reach or exceed expectations, provoking pleasure which directly influences satisfaction (see de Rojas and Camarero, 2008 for further discussion).

Jeong and Lee (2006) measured visitor satisfaction across thirty museums in South Korea using survey scales for: emotional affect, defined as learning, aesthetic experience and enjoyment; overall visitor satisfaction, measured on a 5-point Likert scale from "very unsatisfactory" to "very satisfactory", and fatigue, determined by how exhausted the visitors felt after visiting the museum measured on a 5-point Likert scale from "never exhausted" to "very exhausted". Using factor analysis, the physical characteristics of the museum environment were categorised into three subgroups: the exhibition environment; the ambient environment and; museum size. Exhibition environment turned out to be the main factor that had the biggest effect on satisfaction. The ambient environment had a small, indirect effect on satisfaction. The size of the museum had a slight direct effect on satisfaction, but that was mainly on the mental and physical fatigue felt by visitors.

De Rojas and Camarero (2008) measured the experience, mood and satisfaction of 104 visitors to the Queen Isabel Interpretation Centre in Valladolid, Spain. Visitors were quizzed on their expectations, as captured by a scale constructed from a list of seven items taken from Higgs et al. (2005), and the perceived quality of their experience. The study found a significant relationship between satisfaction and intensity of use.

Packer and Bond (2010) measured the restorative qualities and benefits of the museum environment of four cultural sites across Australia on 307 visitors and 274 local residents through Likert-scale responses for visitors' preferences, motivation, satisfaction, and their probability of returning. Multiple regression analysis was used to examine the relationships between attributes of the environment, aspects of the experience, and the attainment of restorative outcomes. Cognitive experience (gaining new knowledge) was seen as the most satisfying experience by visitors to the museum. There were no significant differences between frequent and infrequent visitors in the kinds of experience they considered most satisfying, but there were between tourists and local visitors.

Brida et al. (2013) measured determinants of satisfaction in the museum of Antioquia in Colombia on a sample of 553 visitors. They found that visitor attributes, such as being a foreigner, being older and having paid for entrance, increased the likelihood of greater satisfaction. Likewise, better signage and treatment by the staff of exhibition rooms positively influenced satisfaction levels.

Del Chiappa et al. (2013) measured the satisfaction of 359 visitors to the Archaeological Museum in Sassari, Italy. The authors applied an ordered logit model using a satisfaction function presented according to museum characteristics, prices, income, time allocated to visiting the museum, individual non-observable characteristics and site non-observable characteristics. The survey presented Likert-scale questions on motivation, satisfaction and loyalty. Key determinants of satisfaction turned out to be quality of experience ratings and the length of the time spent at the museum. Factors such as nationality, gender, and education also acted as significant drivers of visitors' overall level of satisfaction.

Brida et al. (2012), in contrast, surveyed 350 visitors to the Museum of Modern and Contemporary Art in Rovereto, Italy on their commuting experience. They found that travel distance had a negative and highly statistically significant impact on the expected number of visits to the museum: the further the distance from Rovereto, the smaller the probability of repeat visits. Accommodation costs and expenditure on souvenirs at the museum had a significant association with the number of visits too.

Brida et al. (2012) applied the same ITC model to 724 visitors at the Archaeological Ötzi museum, Bolzano. Here, they found that travel costs, food and beverage expenses, and spending for shopping in the city of Bolzano had a marginal positive effect on the repeat visitation rate. Entrance fees had a negative effect, and possibly more so among unemployed and retired visitors with lower levels of income.

A number of studies compare the results from applying CV and RP to the same institution. For example, Martin (1994) determined the appropriate size of museum subsidies in Québec using a CV tax increase payment mechanism to support all museums alongside a travel cost model (TCM), estimating the demand curve through the distribution of visitors to the Musée de la Civilisation by zones of origin and the transport cost from these zones to the museum (following Clawson and Knetsch, 2013). Variables were included to capture the presence of substitute museums and how close the individual lived to the city. CV surveys were collected from 908 Quebec residents.

Martin (1994) found a mean non-use WTP of US\$7.95 (£9.29 present day GBP) per year on a per capita basis to support all Québec museums. TCM analysis of 1,231 museum visitors (use value) found an average of \$8.39 (£10.27 present day GBP) of consumer surplus per Québec visitor (adult and child). Proportional to all museum visitors, total consumer surplus was estimated at \$5.18 million (£6.35 million present day GBP). The total use value of the Musée de la Civilisation was calculated at \$6 million (£7.34 million present day GBP) when combined with \$823,000 (£1 million present day GBP) entrance fee income. Although the CV and RP estimations of WTP were of the same order of magnitude, Noonan et al. (2003) noted that the survey was limited in its ability to accurately elicit WTPs and that the cultural good actually being valued was government spending on museums.

Bedate et al. (2004) measured willingness to travel to four cultural heritage sites in Spain, including the Museum of Burgos. The study classified visitors according to their zone of origin and obtained visits per capita for each of the zones by dividing the number of visitors by the population of that zone. This provided a sample of 49 respondents from bordering zones, 125 from central zones, 101 from peripheral zones and 19 from non-peninsular zones (foreign visitors). WTP ranged from 0.83EU (£0.81 present day GBP) per person for the Iberian Organ Festival to 3.98EU (£3.92 present day GBP) per person for the Museum of Burgos. Previous literature reviews have noted that Bedate et al.'s estimates of use value may be biased upwards "since the survey was sampled at the town's museum, and not all visitors who visit the town may visit the museum" (Eftec, 2005, p. 44).

Boter et al. (2005) compared the relative value of 82 competing museums for a sample of 80,821 museum cardholders. They used a latent class application of a logit model to analyse the impact of the different distances of the population to the museums and differences in cardholders' willingness to travel. Four segments were identified, differing in their sensitivity to distance. Segments were also divided between those where the top 10 museums by willingness to travel were located in the major cities and those where the top 10 museums were located in rural areas. The authors concluded that other amenities available in urban areas may have influenced willingness to travel.

Fonseca and Rebelo (2010) valued willingness to travel to the Alto Douro Wine Region World Heritage Site museum in Portugal. The researchers collected survey data on 372 individuals' residence, the duration of their journey, the number of occupants in the vehicle, and whether the journey was undertaken in a private vehicle, estimating a Poisson model to model the probability of visiting the museum (following Bedate et al., 2004). The authors found that the average travel cost for a visit was 36.40EU (£37.89 present day GBP), with an average of 1.98 visits made.

Armbrecht (2014) compared WTP elicited with open-ended questions to TCM calculated as home distance from site. The study surveyed 583 visitors at the Vara Konserthus and 414 visitors at the Nordic Watercolour Museum in Sweden. Distinguishing between local and non-local visitors to the museum, they reported a mean WTP per person of 32.30EU (£26.93 present day GBP) and 42.30EU (£35.27 present day GBP) respectively. Aggregated WTP (based on mean values and number of visits) was 1.23EU million (£1.03 million present day GBP) for local museum visitors and 4.72EU million for non-locals (£3.94 million present day GBP), giving a total of 5.96EU million (£4.97 million present day GBP). The TCM suggested a consumer surplus of 1.68EU million (£1.1 million present day GBP) for the museum and 552,000EU (£460,000 present day GBP) for the concert hall. Total willingness to pay was calculated as consumer surplus plus travel costs and entrance fee. For the museum, total WTP came out as 5.1EU million (£4.21 million present day GBP), and for the concert hall 1.6EU million (£1.33 million present day GBP). In comparing the results from the CV and TCM, the authors noted that although similar values were found in aggregate, the direct CV yielded much lower figures.

Preferences may also be revealed through other markets, such as museum membership levels. Paswan and Troy (2004) surveyed the membership motivations of 525 museum members. They found a significant positive relationship between membership levels and philanthropy motives and social recognition motives. A significant negative relationship was found between membership levels and children's benefits motives and tangible benefits motives. Hedonic factors (e.g., love of art) were not associated with membership levels either.

### 3.3 Benefits transfer studies

Benefits transfer (BT), the exercise of transposing 'primary' research findings from one site to another, has obvious attractions to funders as an alternative to original, and costly, valuation research. It takes the estimated average WTP values from one study site and applies them to a policy site ('unit estimate transfer'), or 'transfers' the information from the study site to the policy site by using the estimated relationship between WTP and a number of explanatory variables ('function transfer') (Brouwer 2000).

Although widely used in the environmental and health valuation fields, applications of BT to the valuation of cultural and heritage assets are still rare. Arguably this is because sites of cultural significance are considered unique in terms of their characteristics – their architecture, location, historic context, associations and symbolic meaning, or national and international significance, for example, which reduces the validity of transferability (Provins *et al.* 2010). Nevertheless, in cases where the asset and the respective policy change are of a similar nature and significance (e.g. non-iconic historic houses, churches, castles and palaces with similar architecture, exterior and interior decorations; museum collections sharing similar types of objects like paintings, textiles, wooden objects etc.), or importantly where the range of services/benefits provided are similar (leisure, recreation and tourism, education and knowledge, religious and spiritual benefits, and heritage, sense of place and identity, non-use values and so forth), benefits transfer might provide a useful alternative to conducting an original valuation study at each site.

The few existing BT studies of cultural heritage have transferred benefit estimates for a variety of cultural goods, namely for the protection of historic temples in Vietnam and Thailand (Tuan *et al.* 2009), petroglyph images in the US (Ulibarri and Ulibarri, 2009) and aboriginal cultural heritage sites (Rolfe and Windle, 2003). In the UK, Eftec (2005) developed a value transfer study of existing heritage CV studies of Denbigh Townscape, Kennet & Avon Canal, Battersea Park, Lincoln Cathedral, and Sandal Castle. More recently, in 2014, the EU-funded Climate for Culture project conducted a large scale benefits transfer study of a range of cultural institutions in Europe, including museums ([www.climateforculture.eu](http://www.climateforculture.eu)). Focusing on non-iconic heritage and, specifically, the indoor contents of heritage buildings which are arguably largely common across regions (e.g. collections such as paintings, wooden objects, textiles etc.), the authors investigated the reliability of value transfers both within national boundaries (i.e. within same country case studies) and across national boundaries (i.e. different case studies in Europe). They found that heritage conservation values can successfully be predicted via value transfer approaches, with moderate transfer errors. Much more empirical work is needed however to test the transferability of values in the case of museums and galleries.

### 3.4 Wellbeing valuation studies

Fujiwara (2013a) applied the WV approach to Taking Part, a UK cross-sectional survey of 14,000 adult individuals per year from 2005, run by the DCMS, and that asks a wide range of questions on involvement and attitudes concerning arts, culture and sport. Taking Part contains one question related to SWB: "Taking all things together how happy would you say you are?" on a scale of 1-10 (10 = 'extremely happy' and 1 = 'extremely unhappy'). The study applied the two-stage least squares income model (Fujiwara, 2013c) to estimate a compensating surplus associated with visiting museums in one's free time of about £3,200 per year (the amount of money people would in theory give up in order to undertake the activity, related to the concept of willingness to pay). The author noted that the value derived for museum visits far exceeded the value derived from a contingent valuation of other cultural institutions, such as the Bolton Library (2005) which, recall, found a maximum annual WTP of £33.

Fujiwara *et al.* (2014) applied the WV approach to data on arts engagement from Wave 2 of Understanding Society (2010-2011), a nationally representative panel of 40,000 households in the UK. The museums/libraries/heritage sites visit variables in this dataset contained the frequency with which people visited museums, galleries, public libraries and heritage or historical sites in their own time over the previous year. Again using the two-stage least squares income model, Fujiwara *et al.* found that visits to museums was not significantly associated with life satisfaction (though the authors noted the possibility that with larger samples, the coefficients would be estimated with more precision and so museum visits might become significant). The authors did find, however, that engagement with arts and culture more generally was associated with higher wellbeing (valued at £1,084 per person per year), while frequent library use was valued at £1,359 per person per year for those that made use of libraries.

### 3.5 Gaps in the literature

The results of our literature review of cultural valuation in museums and galleries point to three important gaps in the evidence base which the present study sets out to address.

First, comparisons between different valuation methodologies have been confined to a handful of revealed preference and contingent valuation studies (see e.g., Armbrrecht, 2014; Bedate et al., 2004; Boter et al., 2005; Martin, 1994). To date, there have been no comparative studies at all looking at contingent valuation and wellbeing valuation. Second, there have been only a small number of studies using wellbeing valuation in particular to measure economic value in individual cultural institutions – in part, of course, reflecting the novelty of the method (Fujiwara, 2013c; Fujiwara et al., 2014). Third, there are surprisingly few examples of best practice applications of CV in important cultural institutions, based on large samples. The EU Climate for Culture project mentioned above contains a number of large-scale, comprehensive CV studies of cultural institutions, but the results relate to climate change damages only<sup>6</sup>. UK museums, galleries and other cultural institutions would benefit from clear best practice guidance on selecting hypothetical scenarios, payment mechanisms, elicitation mechanisms, and measures to reduce hypothetical bias.

## 4. Case study one: The Natural History Museum

The Natural History Museum (NHM) in London is the third most popular free UK visitor attraction, with over 5.4 million visits in 2013. It was the winner of 'Best of the Best' in the Museums and Heritage Awards 2013. Founded in 1881, it is home to the largest and most important natural history collection in the world, with 80 million specimens, covering all groups of animals, plants, minerals and fossils from across the world, ranging from microscopic cells to dinosaur skeletons.<sup>7</sup>

Entrance to the museum and its main galleries and permanent collections is free. The museum only charges for entry to special temporary exhibitions, such as the popular Wildlife Photographer of the Year. As such, around 60 per cent of the Museum's funding comes from a Government grant.

As well as a world-class visitor attraction, the NHM is also a leading science research centre, employing some 300 scientists who between them published 720 scientific papers in 2013/14. They care for and develop the museum's unique collection, gathered over 400 years, and work to conserve the richness and diversity of the natural world with projects in more than 68 countries. In 2009, the NHM in London opened the Darwin Centre, giving the public the opportunity to observe the work of its scientists at first hand and participate in scientist-led events and activities. The Centre is also a research facility for the Museum's scientists.

The NHM also has a sister museum in Tring, Hertfordshire, which attracted 154,000 visitors in 2013/14 to its changing programme of exhibitions and activities. It also played host to scientific researchers visiting the NHM Tring's extensive bird collections.

One of the UK's leading cultural institutions, producing a range of market and non-market benefits for the public and in receipt of significant public funding, the Natural History Museum is a compelling case study to explore the applicability of economic valuation techniques endorsed by HM Treasury's Green Book.

## 5. Case study one: Survey methodology

We conducted two surveys, each combining contingent valuation with wellbeing measurement. The first survey was a face-to-face exit survey of visitors to the NHM in London (N=616). The second was an online survey of the UK population (N=1,000). Visitor and general population (online) surveys were screened to exclude UK non-residents and those under sixteen years old. For the online survey, quotas for gender, age, and region were set using national UK averages from the Office for National Statistics' Annual Population Survey. On-site interviewers followed a protocol of randomly selecting one in every three people exiting the museum. We then applied NHM visitor weights for gender and age (assembled between October-December 2013) to the visitor survey data to correct for slight over-representation of older people and a slight gender imbalance (see weighting table, Annex D6).

<sup>6</sup> Our study draws on the lessons learned from this large scale project, however (as one of the Principal Investigators of this research is also co-author of this study).

<sup>7</sup> Natural History Museum Annual Report and Accounts 2013/14 <http://nhm.ac.uk/resources-rx/files/annual-report--accounts-2013-14-132416.pdf>

The survey was divided into four sections. The first section contained background questions on the size of respondents' visitor groups, number of children, use and enjoyment of the Natural History Museum, and whether they had made a donation during their visit. The visitor survey included specific questions on respondents' recent experiences in the NHM in London, which parts of the museum they had most enjoyed, and their reasons for visiting. The online survey asked whether respondents had heard of, visited, and enjoyed the NHM at London and at Tring, and their likelihood of visiting again. Both surveys asked identical questions about membership of heritage, conservation or environmental organisations, and their use of other museums and cultural sites in the last twelve months. Respondents in both surveys were asked an identical prioritization question on the allocation of government funding, and a set of questions on cultural engagement while growing up, developed from the Taking Part survey. A set of statements about the value of the NHM were presented on a five-point Likert scale (1, strongly disagree; 5, strongly agree). These statements varied across surveys to reflect the difference in use and non-use value elicited, with the exception of two statements on value of the NHM for future generations and value for those people who do not visit, which were identical across both surveys.

In the second section, both sets of respondents were asked the following three SWB questions:

1. *Overall, how satisfied are you with your life nowadays? ('Life satisfaction')*
2. *How happy were you feeling in the last hour or so? ('Happiness')*
3. *How worthwhile did what you were doing in the last hour or so feel? ('Purpose')*

People responded on a scale of 0-10 where 0 = 'Not at all' and 10 = 'Completely'. In the visitor survey, respondents were asked which museum activity they had been doing in the previous hour. In the online survey, we asked what activities the respondent had been doing in the previous hour, developed from the Mappiness iPhone Experience Sampling Method (ESM) (MacKerron and Mourato, 2013).

In both surveys we asked respondents if they thought they would be happier if they could visit the Natural History Museum more often. Specifically, a life satisfaction vignette scenario was presented to all respondents. The vignette asked respondents to:

"Take your life as it currently is, but now imagine that you have visited the Natural History Museum [once every [frequency] months (in other words, about [frequency] over the year)] alone or with your family/friends. Assume that you usually spent one to two hours there per visit and vary the activities each time". Overall how satisfied would you be with your life?"

We randomised the visit frequency between once every six, four, three, two and one month(s) in order to derive unbiased estimates of the impact of NHM visits on life satisfaction.

The third section presented respondents in both surveys with identical information on the Natural History Museum, in terms of the services and visitor attractions it offered, and the conservation and research work it undertakes. The online survey provided information on the NHM's site at Tring. Respondents were asked about their familiarity with the information before this survey on a five-point Likert scale (1, not at all familiar; 5, extremely familiar).

Willingness to pay scenarios, payment mechanism, and elicitation method differed between the visitor and online survey. Below we describe the key components of the design of the CV and the SWB components of the surveys and their methods of analysis.

The fourth and final section of the surveys asked a set of standard socio-demographic questions for use in the CV and SWB analysis.

## 5.1 Contingent valuation

The visitor survey contingent valuation scenarios were designed to uncover the following values:

- (A) the value of accessing the NHM (via an entrance fee)
- (B) the value of the conservation and research services provided by the NHM (via a donation, elicited as a top-up to an entrance fee)

The general population survey contingent valuation scenarios were designed to uncover:

- (C) the value of the conservation and research services provided by the NHM (via an annual donation)

The 'hybrid' contingent-wellbeing valuation question was asked of both visitors and the general population, designed to uncover:

- (D) the value of avoiding the closure of the NHM for one year (via a one-off cash compensation).

In terms of the Total Economic Value framework: (A) is a direct use value; (B) is a non-use value with possible elements of option value (in that the conservation and research work of the NHM may affect the collections and the way they are curated and presented); (C) is similar to (B), i.e. a non-use value with possible elements of option value, and (D) is likely to be mostly a use-value and an option value, as closure would prevent access and future access but not the on-going research and conservation.

The following section provides further details about these scenarios.

### 5.1.1 Contingent scenarios design

As noted above, the visitor survey explored use values related to the NHM using stated preference methods. Respondents were asked to state the maximum they would be willing to pay as an entrance fee to the NHM and how much compensation they would require if the NHM were to close for a period of time (and hence they could not visit). The compensation question was also designed to be used in the wellbeing analysis.

In the visitor survey we asked two contingent valuation questions.

The first, relating to use values, asked questions about the value of a visit to the Natural History Museum (see Box A1). We provided two paragraphs of information on the visitor services the Natural History Museum offers, including recreation, education, inspiration, and asked respondents about their familiarity with the information.

The payment vehicle used was an entrance fee, asked only to those over 16 years of age. Respondents were given the choice of whether they would require an individual ticket or a family ticket, and values in the payment card presented accordingly. We laid out a hypothetical scenario where museums in the UK were considering charging for access due to restrictions in government funding. In CV surveys, in order to elicit welfare-consistent values it is vital that the hypothetical scenario be believable, and that it justifies introduction of the hypothetical payment mechanism (in this case an entry fee). Interviewers were instructed to stress that this was a hypothetical scenario. We presented respondents with a payment card offering them a range of monetary amounts they could choose from. This provided respondents with a visual aid and helped remove starting point bias (Bateman et al., 2002; Maddison and Foster, 2003; Maddison and Mourato, 2001).

#### Box A1. Visitor survey entrance fee (use) willingness to pay question

The current financial crisis has meant that the Natural History Museum has suffered cuts in Government funding, while having to cope with record numbers of visitors and associated increases in maintenance and operating costs. In the unlikely event that Government funding ceases to be provided, the Museum would need to charge an entrance fee to raise enough money to support its activities and secure its long-time future.

For the next set of questions, please imagine a situation where Government funding for museums has been cut and museums start charging an entrance fee. In this situation, please think about how much a visit to the Natural History Museum would be worth to you and your family, if anything.

*If the Museum were no longer free, what is the maximum you would be willing to pay, **as an entrance fee, [just for yourself/ for a family ticket]**, to visit the Natural History Museum?*

In answering this question, please focus solely on how much a visit to the Natural History Museum is worth to you and how much you enjoyed the visit. Please do not consider the value to you of other activities that the museum carries out. In this question, we are just interested in how much benefit you get from a visit. Also please be realistic – consider your household budget and remember that there may be other things you would like to spend your money on, including visiting other museums that may also charge for entry.

*If Yes/Maybe: Would you be prepared to pay an entrance fee to visit the Museum, even if only a small amount?*

In answering this question, please focus solely on how much a visit to the Natural History Museum is worth to you and your family and how much you enjoyed the visit. Please do not consider other activities that the museum carries out. In this question, we are just interested in how much benefit you get from a visit. Also please be realistic – consider your household budget and remember that there may be other things you would like to spend your money on, including visiting other museums that may also charge for entry.

WTP values were elicited using a payment card with values ranging from £0 to £50 (£100 in the case of a family ticket).

As discussed earlier, the second contingent valuation question was intended to explore mostly non-use values, asking respondents to value the work performed by the Natural History Museum in caring for a national collection, conservation work, and scientific research. We provided respondents with a paragraph of information on the conservation and research activities of the museum, and asked their familiarity with the information. The payment vehicle was a donation (per family) on top of the entrance fee they had already indicated they would pay (see Box A2). Respondents were reminded of this amount in the second contingent valuation question.

### **Box A2. Visitor survey donation (non-use) willingness to pay question**

In order to help fund all its crucial conservation work and scientific research, the Museum could offer an admission ticket which includes a voluntary donation on top of the entry price. This is currently already done for entry tickets to the special temporary exhibitions (such as the Wildlife Photographer of the Year exhibitions). Many other institutions also offer entry tickets with a small donation added for conservation work (e.g. London Zoo, National Trust properties). Note that this would be totally voluntary and visitors could readily opt to pay the entrance fee without the donation.

*Please think about how much the conservation and research work carried out by the Natural History Museum is worth to you, if anything. If the Museum were no longer free, you said you would be prepared to pay an entrance fee of £X to visit [INSERT ENTRANCE FEE VALUE]. What is the maximum you would be willing to pay, **as a donation on top of this entrance fee**, towards the conservation and research work carried out by the Museum?*

WTP values were elicited using a payment card with values ranging from £0 to £15.

The online general population survey was intended to explore mostly non-use and option values. We asked respondents to imagine that the NHM was considering launching an appeal to fund the work involved in 'caring for a national collection', including its conservation work, and scientific research programme, necessary due to restrictions in Government funding because of the current financial situation (see Box A3). We provided respondents with information about the conservation, research, and scientific work of the NHM in London and Tring, using illustrative photographs, and elicited how familiar they had been with the information beforehand. We used the payment vehicle of an annual donation (per family).

### **Box A3. Online survey donation (non-use) willingness to pay question**

Entrance to the Natural History Museum (in London and Tring) and its main galleries and permanent collections is free. The museum only charges for entry to special temporary exhibitions. As such, around 60 per cent of the Museum's funding comes from a Government grant.

The current financial crisis has meant that the Natural History Museum has suffered cuts in Government funding.

**If Government funding ceases to be provided, the Museum would need to raise enough money in donations to secure its long-time future and support its activities, including its crucial conservation work and scientific research.**

For the next set of questions, please imagine a situation where Government funding for museums has been cut. In the unlikely event that Government funding ceased, please think about how much the collections and the conservation and research work of the Natural History Museum (in London and Tring) would be worth to you and your family, if anything.

**Would you be prepared to pay a donation, even if only a small amount, to secure the future of the Natural History Museum's collections, conservation work, and scientific research programme?**

**If Yes/Maybe: Please think about how much the continued existence of the Natural History Museum (in London and Tring) is worth to you, if anything. What is the maximum you would be willing to pay, as a donation, per year, to secure the future of the Museum's collections, conservation work, and scientific research programme?**

Studies have shown that many people answering surveys such as this one, say they are willing to pay more in donations than they would actually pay in reality. So please think about this question as if it were a real decision and you were actually making a donation for real. Please do not agree to pay an amount if you think you cannot afford it, if you feel you have paid enough already, or have other things to spend your money on, and other ways to fund conservation and research. Also, the question is just about the work of the Natural History Museum and not about other museums.

WTP values were elicited using a payment card with values ranging from £0 to £150.

In both surveys, respondents were asked a certainty question on the donation amount that they had stated. Respondents were also asked to give reasons for their willingness, or not, to pay.

In designing the contingent valuation scenarios (for both the visitor and online surveys) we implemented best practice to attempt to deal with the known biases in CV, as discussed in Section 2.2.1.<sup>8</sup> These related to question wording and ordering, payment format, payment vehicle, prompts and use of images (Bateman et al., 2002). Respondents were 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 (Champ and Bishop, 2001; Cummings and Taylor, 1999). Certainty questions were asked after the two contingent valuation questions (Bedate et al., 2009; Champ and Bishop, 2001), and we collected follow-up information on the reasons that they were or were not willing to pay the amount stated (Bateman et al., 2002).

We ran a randomised experiment within the online survey to test best practice in reducing hypothetical bias in contingent valuation methods. We set up a four-branch experiment for this purpose:

1. a cheap talk script informing respondents that “studies have shown that many people answering surveys such as this one, say they are willing to pay more in donations than they would actually pay in reality”;
2. an oath script asking respondents to “promise to answer the questions that will follow as truthfully as possible”;
3. cheap talk and oath scripts in combination;
4. neither oath script nor cheap talk.

In all cases, following best practice, respondents were reminded to be realistic, reminding them of the household budgetary constraints, and the existence of alternative cultural institutions that they might wish to spend their money on (Champ and Bishop, 2001; Cummings and Taylor, 1999).

### 5.1.2 Contingent valuation analysis

As described above, the visitor and general population surveys provided three separate WTP and one WTA payment mechanisms.

- A combined individual and family entrance fee. This was created by adding WTP for entry with an individual ticket value (n=413) to family ticket values (n=113), calculated proportional to the size of the visiting group (family and adults). This provided a use value for the full sample of NHM visitors (n=616). (Note that No responses in the binary choice WTP question (n=90) were assigned in full to both the individual and family ticket variables, accounting for the difference between total use sample and aggregate observations).
- A donation on top of any entrance fee, asked of the full sample of n=616 visitors;
- An annual donation towards the continued existence of the Natural History Museums in London and asked in the online general population survey (n=996).
- A cash compensation to all respondents across both surveys (n=1,260). The analysis of WTA is covered in detail in Section 5.2.2.2.

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 in between the amount they choose and the next amount on the payment card.

We calculated non-parametric mean and median WTP and WTA from the mid-point between the amount chosen on the card and the next amount up. The latter is theoretically consistent with the statistical theory for calculating WTP from interval data (Cameron and Huppert, 1989). All No responses were coded as £0 bids. We removed outliers above £500 as protest votes (n=2) and coded unanswered open-space responses as missing.

<sup>8</sup> In doing so we improve on the few existing applications there have been of CV methods in the UK’s cultural sector. We will also discuss related findings from the on-going EU project Climate for Culture that has conducted contingent valuation studies at ten heritage sites in Europe, including Ham House, Knole House and St Joseph’s and the English Martyrs Roman Catholic Church in the UK.

The following regression model was used as the base for all the WTP analyses:

$$WTP_i = \alpha + \beta_1 X_i + \varepsilon_i \quad (5)$$

where  $WTP_i$  is the amount the individual  $i$  has stated they are willing to pay,  $X_i$  is the deterministic factor and  $\varepsilon$  is the error term containing unobserved factors that determine willingness to pay. In  $X_i$  we control for the observed determinants of willingness to pay (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.

We estimated three variants of equation (5) for each of the three WTP measures, meaning nine models in total.

- Model 1: Equation (5) with socio-demographic variables
- Model 2: Model 1 with variables capturing experience at the museum
- Model 3: Model 2 with attitudes, opinions, and proxy variables for cultural engagement

We tested the distribution of residuals for heteroskedasticity using robust standard errors and for normality using kernel density estimates (Annex D7).

We applied a number of tests on the validity of our results. For instance, we tested for differences in WTP between those who indicated that they were very certain with the answers they gave, and those that were either uncertain or somewhat certain using independent sample  $t$ -tests. Specifically, we estimated the following model using OLS:

$$WTP_i = \alpha + \beta_1 Cert_i + \beta_2 X_i + \varepsilon_i \quad (6)$$

where  $WTP_i$  is the amount the individual  $i$  has stated they are willing to pay,  $Cert_i$  is the individual's stated certainty to pay that value  $X_i$  controls for determinants of WTP in Model 1,  $\alpha$  is the deterministic factor and  $\varepsilon$  is the error term containing unobserved factors that influence WTP.

We explored possible protest bids by analysing the reasons given by respondents for being willing or not willing to pay an entrance fee in the case of the visitor survey, and donation in the case of the online general population survey (Annex D3). Since in the event the estimated number of protests was small (see Section 6.4.2.1), all responses were retained in the analysis.

We tested for differences in stated WTP between individuals randomly shown one of the four experimental cheap-talk/oath branches in the online survey using the following model:

$$WTP_i = \alpha + \beta_1 CT_i + \beta_2 OA_i + \beta_3 CTOA_i + \beta_2 X_i + \varepsilon_i \quad (7)$$

where  $WTP_i$  was the amount the individual  $i$  had stated they were willing to pay,  $CT_i$  is a dummy variable that denotes whether an individual was presented with only the cheap talk script,  $OA_i$  is a dummy variable marking whether an individual was presented with an oath script only,  $CTOA_i$  captures when an individual was presented with the cheap talk and oath script together and,  $X_i$  controls for determinants of WTP in Model 1,  $\alpha$  is the deterministic factor and  $\varepsilon$  is the error term containing unobserved determinants.

This allowed us to quantify the additional effect of an oath instrument or cheap talk script on stated WTP, and to test the interactive effect of the oath and cheap talk together – hence guiding us on how to minimise hypothetical bias (studies to date have focused on either oath or cheap talk not both; see e.g., Carlsson et al., 2013; de-Magistris and Pascucci, 2014; Stevens et al., 2013).

## 5.2 Subjective wellbeing

### 5.2.1 Subjective wellbeing design

Both surveys asked respondents questions about their SWB and a set of socio-demographic variables. Variation in NHM visit frequency across the full sample was used to assess the impact of NHM visits and engagement on life satisfaction after controlling for a wide range of factors.

We also collected data on the individual's main activity over the past hour to assess the drivers of momentary wellbeing, as captured by their happiness and purpose. This approach broadly followed an Experience Sampling Method (ESM) design. In ESM, participants are asked to stop at certain times and to make notes of their experience in real time. They report how they feel – their wellbeing – and what activity they are doing. This can be done through a journal diary, using a personal data assistant (PDA) or a smartphone app. The main difference with our surveys is that we ask people to reflect on a very recent period in time (the previous

hour) rather than the present, which may help to lower the risk that data collection will disturb people during their activities and interfere with their actual experiences – a well-known problem in ESM (Kahneman et al., 2004).

In addition, a vignette was used to further assess the association between NHM visits and life satisfaction. Respondents were asked to imagine a situation where they were able to visit the NHM more frequently and were asked what their level of life satisfaction would be, holding all other factors in their lives constant. An important consideration in using the vignette was the fact that the NHM has relatively few repeat visitors,<sup>9</sup> so variation in NHM visits frequency in a sample study turns out to be low. Through the vignette we are able to alter visit frequency to produce more variation. A full discussion of the rationale, benefits and challenges of employing vignettes in the NHM case study is provided in Section 5.2.2.2.

Results for associations with life satisfaction can be used to derive monetary values using the wellbeing valuation approach as discussed in more detail in Section 5.2.2.2. The questions on happiness and purpose in the moment should be viewed as providing information to support the life satisfaction results, and tell us at a more fine-grained level the relationship between specific museum activities and SWB.

The visitor and online surveys for the NHM case study can be found in Annex A.

### 5.2.2. Subjective wellbeing analysis

There are two aims of the SWB analysis:

- (i) To assess the impacts of NHM activities and visits on momentary wellbeing.
- (ii) To estimate the values that people place on engagement with the NHM using the wellbeing valuation approach and to compare these values with values estimated using CV. This analysis uses life satisfaction – a non-momentary measure of wellbeing.

The analysis on momentary wellbeing should be understood as complementing and informing the findings of the wellbeing valuation.

#### 5.2.2.1. Momentary wellbeing

We assess the relationship between different cultural and non-cultural activities and momentary SWB, measured as happiness and purpose, over a short period of time (the last hour) using the following econometric model:

$$SWB_i = \alpha + \beta_1 A_i + \beta_2 X_i + \varepsilon_i \quad (8)$$

where  $SWB_i$  is individual  $i$ 's subjective wellbeing (happiness or purpose) over the previous hour;  $A_i$  is a list of activities that the individual undertook during this period (where working/studying is used as the reference case in the models);  $X_i$  is a vector of control variables and  $\varepsilon_i$  is an error term. We use the following variables to control for selection into certain activities: age, gender, health, income, marital status, education, parental status, employment status, social networks and current life satisfaction. All these variables have been shown in the literature to correlate with happiness, and may also correlate with the propensity to select certain activities.

As with any study using non-experimental data – that is, where treatment or activities have not been strictly randomly assigned – there is a residual risk of selection bias and this should be noted when interpreting the results. However, the fact that we have used an extensive set of control variables and that we relate activities to SWB within a very specific time frame should add more confidence to our results.

All SWB models in this paper are estimated using ordinary least squares (OLS) regression analysis. This assumes that the SWB reporting scale (0 to 10) is cardinal. Ferrer-i-Carbonell and Frijters (2004) show that it makes little difference in wellbeing models whether one assumes cardinality or ordinality in the wellbeing variable. As a consequence, much of the literature now uses OLS rather than ordered probit models because of the ease of interpretation of OLS results – in particular, OLS estimates imply impact estimates in terms of changes in absolute levels of SWB. We use heteroskedasticity-robust standard errors in all models. This is in line with best-practice in the wellbeing literature in order to address the common observation of heteroskedasticity in large sample data.

<sup>9</sup> Reference numbers from the NHM.

### 5.2.2.2. Wellbeing valuation

The wellbeing valuation (WV) approach derives monetary values for non-market goods and services – like health, culture and education – by estimating the amount of money required to keep individuals just as happy or satisfied with life in the absence of the good (i.e., to keep their wellbeing constant). This is the fundamental idea that underlies welfare economic theories of value such as compensating and equivalent surplus.

Here, we derive estimates of compensating surplus (CS) for NHM engagement. CS is the amount of money, paid or received, that will leave the individual in his initial welfare position following a change in the (level of a) good/service. CS is the most widely used measure of value in CBA and is akin to the notion of willingness to pay.

We employ three different approaches to WV, two of which, as far as we are aware, have not been used previously.

#### (i) Standard wellbeing valuation approach

First, we estimate the value of NHM engagement using the naturally occurring variation in NHM visit frequency across the sample population. In this sub-section, we provide a concise description of the standard WV method; a full description of the econometric methodology used in WV can be found in Annex C.

The standard WV approach uses the outputs from a life satisfaction model of the following type:

$$LS_i = \alpha + \beta_1 M_i + \beta_2 NHM_i + \beta_3 X_i + \varepsilon_i \quad (9)$$

where  $LS_i$  is life satisfaction for individual  $i$ ;  $M_i$  is income;  $NHM_i$  is a variable proxying engagement with the NHM; and  $X_i$  is a vector of other determinants of life satisfaction. In the analysis, NHM engagement is represented by a variable denoting whether the individual has visited the NHM over the past year. As in the momentary wellbeing models,  $X_i$  include the main determinants of wellbeing (see Fujiwara and Campbell, 2011 for a review): age, gender, health, income, marital status, education, parental status, employment status, social networks.

In addition, we control for whether the survey was taken online or in person to account for possible survey mode effects, and any unobserved differences present when the surveys were completed, such as in the weather. We also control for a set of factors that make selection into museum engagement more likely. These are variables that have been found to be important determinants of engagement in museums and cultural activities in previous studies: whether or not an individual is a member of the NHM; whether they were taken to museums (or arts galleries) as a child; whether they took any classes or training in the arts; whether they live near a museum or gallery, and whether friends and family often go to museums or galleries. All these variables might also plausibly correlate with subjective wellbeing and therefore their exclusion from the model risks giving rise to omitted variable bias.

The economic value of NHM engagement can, in principle, be assessed by using the estimates of the coefficients for income and NHM (respectively  $\beta_1$  and  $\beta_2$ ).  $\beta_1$  represents our estimate of the effect of income on life satisfaction and  $\beta_2$  represents our estimate of the effect of NHM visits on life satisfaction. The ratio of these two figures is an estimate of the value (CS) of NHM engagement because it shows how much extra income is required to have the same impact on life satisfaction as NHM engagement. Put another way, it is the amount of money required to keep individuals just as satisfied with life in the absence of visiting the NHM (i.e., to keep their wellbeing constant).

$$\text{Value of NHM engagement} = \beta_2 / \beta_1 \quad (10)$$

This is a heuristic description of the approach we follow. In practice we use a more involved method for estimating  $\beta_1$ , which adopts a non-linear structure and which controls for selection bias, though the general principle of comparing the ratio of the two coefficients still applies. The technical background and details of the statistical methodology employed in the WV analysis are set out in Annex C.

The analysis, as with most studies in this area, is necessarily based on observational datasets (i.e., where people have not been assigned to different conditions in a controlled experimental setting). Thus cause and effect relationships are *approximated* using statistical methods, as causation cannot be directly inferred. This is because there are likely to be a number of unobserved factors correlated with museum attendance that differ initially between the different groups and which influence the outcomes we observe. For example, it might be that museum-goers tend to be of a more extrovert character and this personality trait may make them more likely to visit museums as well as being happier. Causation can only be fully inferred in our framework if *all* of these confounding differences are controlled for.

Notwithstanding this difficulty, we follow best practice in WV by controlling for all of the main determinants of wellbeing in the regression analysis. Multiple regression analysis of the type employed here has been used extensively in the academic and policy evaluation literatures, though definitive statements about causality can only be made in a controlled experimental setting – a priority for future wellbeing research.

### (ii) Hybrid contingent-wellbeing valuation

In this approach, we ask respondents directly how much monetary compensation they would require if they were not able to visit the NHM for a period of time (one year) such that their life satisfaction would remain unaffected. Note that this is essentially a contingent valuation study using a WTA elicitation format. A similar approach to the one adopted here was previously described in Bateman et al.'s (2002) manual, although it has not been widely used since. Since the approach makes explicit mention of 'wellbeing' (or life satisfaction in this case) however (see below), we henceforth describe it as a 'hybrid contingent-wellbeing valuation' approach.

We use the following question (Box A4):

#### Box A4: Hybrid contingent-wellbeing valuation question

*For this next question, please imagine that the Natural History Museum had to **close to the public for one year** for vital maintenance work. No one would be able to visit any parts of the Museum during this period. **Other museums would remain open as usual**. Now don't worry, there are no plans for the museum to close! But we would like you to think about what your life would be like if it did close for one year. How would the closure affect you level of life satisfaction?*

- The closure would have very little effect on my life satisfaction*
- The closure would reduce my life satisfaction*
- The closure would increase my life satisfaction*

*If the respondent selects option 2 then,*

*Now imagine the following situation. Suppose that in order to compensate you for not being able to visit the Natural History Museum during one year, you were given a cash compensation. How much money would you have to receive, as a one-off payment, to give you **the same life satisfaction** that you have now (**not better nor worse, but just the same**) during this period until the Museum re-opened? Think about this for a moment please.*

WTA values were elicited using a payment card with values ranging from £0 to £150.

This follows the theoretical approach underlying WV, but the difference is that equation (10) is, in effect, internalised by the respondent. That is, instead of estimating the elements of equation (10) ( $\beta_1$  and  $\beta_2$ ) through data on people's actual visit frequency and life satisfaction we pass on that task to the individual who needs to think about how her life satisfaction is impacted on by a restriction on NHM visits (akin to  $\beta_2$ ) and how much money would compensate (offset) this effect (akin to  $\beta_1$ ).

This part of the study is similar in some respects to the work by Lau et al (2013), who asked survey participants in the UK and Hong Kong their WTP to re-create the experience of feeling [a certain type of mood] for one hour. The moods included happiness, love, fear, sadness and so on (where for negative moods respondents were asked WTP to avoid the mood). As in our study, people were being asked to equate a specific feeling of wellbeing to a monetary figure.

As discussed in Section 2.1.1, willingness to accept (WTA) compensation questions are sometimes used in CV, but – despite the Bateman et al. (2002) example – normally respondents are not asked for compensation in terms of wellbeing/life satisfaction impacts – they are typically just asked for compensation *per se*. Moreover, compensation was only offered in our case to those that expressed a negative wellbeing effect in the first place. In this sense, the approach we adopt brings together elements of CV and WV, using the WTA structure from CV with the specific theory from WV. Our hypothesis is that reminding people that compensation should only *just* compensate them for an impact on wellbeing (and crucially only for those that said their wellbeing would be negatively affected in the first place) will help to reduce upward bias in WTA values (see discussion below). We use life satisfaction in this case since this is a common measure of wellbeing used in surveys and to also ensure that the results can be compared against the results of the standard WV approach. The hybrid contingent-wellbeing valuation study will provide values per visit based on life satisfaction, whereby the overall WTA value for the year will be divided by the average number of visits in a year. This way, the results can be used to complement and better understand the results from the standard WV study.

One other interesting avenue we explore in our study is how close the WTA values replicate the WTP values for entry/visits. As discussed in Section 2.1.1., any stated preference can be presented in terms of WTP or in terms of WTA. Theory suggests that if anything WTA for most goods may exceed WTP by a few percentage points due to the fact that WTP is constrained by an individual's income (Sugden, 2005). Numerous papers have found, however, that stated WTA is often way above stated WTP for the same good (Horowitz and McConnell, 2002; Shogren et al., 1994). Sugden (2005) argues that the most credible explanations for this relate to the psychological arguments concerning loss aversion and its derivative, the endowment effect, whereby people ascribe more value to things merely because they own them (Kahneman and Tversky, 1979; Loewenstein and Adler, 1995).

Some authors argue that the appropriate formulation depends on property rights (Carson et al., 2001). That is, if the respondent is being asked to give up an entitlement, the WTA formulation is appropriate (Carson, 2000). Others argue that the WTP formulation should always be used (Arrow and Solow, 1993). The WTP-WTA disparity may also, to some extent, be a product of informational constraints and inexperience. For example, List (2003) found that the behaviour of more experienced traders (in a number of different real markets) shows no signs of an endowment effect.

Our study is the first to assess whether constraining the WTA scenario by setting it explicitly in the context of changes in life satisfaction produces reasonable WTA values relative to WTP. The hypothesised mechanism for this is that respondents are directed explicitly to think about the WTA question within the framework of economic theory (i.e., compensating a welfare change). We note that our definition of welfare here (i.e., life satisfaction) differs from the standard definition of preference utility in economics, but we would still hypothesize an effect due simply to the explicit reminder to keep welfare constant.

This element of the study is not free of statistical problems, however. Although we do not estimate  $\beta_1$  and  $\beta_2$  statistically, biases will arise if people are unable to predict the impact of museum closure on life satisfaction. Indeed, a number of experiments have shown that people are unable to do this with accuracy (Kahneman, 2000). Kahneman and Snell (1992), for example, report that people find it very hard to predict how much pleasure they will derive from consuming even everyday goods such as music, yoghurt and ice cream. Gilbert (2007) and Kahneman and Snell (1992) attribute these types of findings to a 'presentism heuristic': people project current tastes and desires on to their predicted future tastes and preferences. Clearly though, preference-based valuation methods such as CV will also suffer from these prediction biases.

### (iii) Vignette-based wellbeing valuation

In our third approach to WV, we use a vignette to overcome two possible problems with standard WV. The latter relies on actual variation in NHM visits in order to estimate values. Repeat visits to the NHM are proportionately small and so this creates problems for WV in a sample study. The second problem with standard WV, as discussed above, is that it is difficult to establish causality with any confidence absent an experimental design. That is, that any positive relationship we detect between life satisfaction and NHM visits could be driven by third factors that are omitted or imperfectly captured in the model.

A vignette may, at least in principle, address both problems by presenting individuals with a hypothetical life or event and asking them questions about this case. Anchoring vignettes are brief descriptions of hypothetical situations measuring a single concept, with responses rated on a Likert-type scale (King et al., 2004). Respondents may be questioned on each concept with both self-assessment items and hypothetical items that represent the range of possible levels of the variable being measured.

Anchoring vignettes are increasingly used in wellbeing and health research. For example, Kapteyn et al. (2011) compare self-reported satisfaction with incomes in the Netherlands and the US. They use vignettes to anchor the effect of cultural differences in responses. The vignette question they use is,

*'(Name) is married and has two children; the total after tax household income of his/her family is (INCOME). How satisfied do you think (Name) is with the total income of (his/her) household?'*

Kristensen et al (2006) use vignettes in a similar fashion to Kapteyn et al. (2011) in job satisfaction questions. Angelini et al. (2012) use vignettes to control for cross-country differences in responses to life satisfaction questions across ten European countries. Specifically, they use the following two vignette designs which ask respondents what they think the life satisfaction of a hypothetical person is:

*'John is 63 years old. His wife died 2 years ago and he still spends a lot of time thinking about her. He has four children and ten grandchildren who visit him regularly. John can make ends meet but has no money for extras such as expensive gifts to his grandchildren. He has had to stop working recently due to heart problems. He gets tired easily. Otherwise, he has no serious health conditions. How satisfied with his life do you think John is?'*

*'Carry is 72 years old and a widow. Her total after tax income is about 1,100 per month. She owns the house she lives in and has a large circle of friends. She plays bridge twice a week and goes on vacation regularly with some friends. Lately, she has been suffering from arthritis, which makes working in the house and garden painful. How satisfied with her life do you think Carry is?'*

Our vignette study follows on from these approaches using an anchoring vignette design to determine the effect of engaging with the NHM on life satisfaction. This is achieved by presenting survey respondents with a varying number of visits to the museum by a hypothetical individual and then asking what they think that individual's life satisfaction would be.

The simplest approach to take would be to present two hypothetical individuals with varying levels of NHM engagement (but other things being broadly similar) and asking the respondents to rate the vignettes' life satisfaction scores. There are, however, two potential problems with this approach.

First, as Krosnick (1991) identifies, the respondent has to give answers based on imperfect information and/or an imperfect understanding of the vignettes' lives. Although in the literature respondents have generally been able to provide Likert-scale type responses based on a small amount of information, or have managed to make a choice over which life they would prefer, that is not to say that these responses were accurate.

Second, presenting vignettes together focuses the respondent's mind on the differing factor (here cultural engagement), which leads to problems related to the so-called focusing illusion effect, whereby a respondent pays too much attention to the issue being highlighted in a way that would not be so under more ordinary conditions (Kahneman et al., 2006). This would likely lead to an overstatement of the importance of cultural engagement for life satisfaction, and is supported by Hsee and Zhang's (2004) research on choice under single or joint evaluation modes, which shows that people attach higher value to things that differ under joint evaluation. One possible solution is to vary a number of other aspects of the vignettes' lives to reduce the focusing illusion – as is standard in choice modelling experiments where similar focusing illusion biases can arise – but this requires statistical modelling that generally relies on large sample sizes (and larger than the ones in our study).

The approach we adopt therefore is to use what we call a *first person vignette* (FPV). An FPV takes the hypothetical element of a regular vignette and adds to it the current situation of the individual/respondent. It asks about the life satisfaction score the respondent would attribute to her life under some hypothetical scenario, where some element of her life has changed (hypothetically). The vignette question we use is:

'Take your life as it currently is, but now imagine that you have visited the Natural History Museum in London or Tring [once every two months] (in other words, about [six times] over the year) alone or with your family/friends for the past year or so. Assume that you usually spent one to two hours there per visit and vary the activities each time. Overall how satisfied would you be with your life?'

Not at all												Completely
0	1	2	3	4	5	6	7	8	9	10		

We vary the visit frequency randomly across the sample with a range of [once every SIX months]/[once every FOUR months]/[once every THREE months]/[once every TWO months]/[once every month].

Van Praag et al. (2003) find that third person ratings of SWB are accurate for family members and close friends (for whom respondents have a lot of background information). Hence, we hypothesise that given that people are being asked essentially about their *own* lives, albeit with a minor change in circumstances, they should find it reasonably straightforward to respond to the SWB question as in third party reports, and certainly more easily than for an unfamiliar subject. Consistent with this, results from our pre-testing of the survey with visitors and an online population suggested that respondents found the vignette question generally easy to answer. Only 10% of 65 test respondents found the vignette question difficult to make sense of and 97% said they had no problem with the clarity of language in the vignette.

By adopting the FPV research design we aim to reduce (i) informational problems related to not knowing all aspects of the vignette subject's life, and (ii) focusing illusion bias, since respondents only focus on one individual life – their own (although focusing illusion cannot be totally eradicated since respondents naturally compare the life described in the vignette with their own lives) where the difference in NHM visit frequency would get highlighted.

As in the hybrid contingent-wellbeing valuation, the vignette study provides an alternative approach to valuation using the WV method as a base. But this time, the estimation of the effect of NHM visits ( $\beta_2$ ) is internalised by the respondent; we pass on the task to the individual who needs to think about how her life satisfaction is impacted on by changes in NHM visit frequency. This still leaves the impact of income on life satisfaction ( $\beta_1$ ) to be estimated, however, and we do this separately using the WV

methods set out in Annex C. That is, the vignette approach risks the same problems as does the hybrid contingent-wellbeing valuation, as people are asked to predict impacts on life satisfaction.

The vignette study is used to derive estimates of the impact of NHM visits on life satisfaction and, given the estimated impact of income on life satisfaction, the value of these visits. Specifically, each survey respondent has (i) an actual level of life satisfaction associated with an actual level of NHM visits, and (ii) a hypothetical level of life satisfaction associated with a hypothetical level of NHM visits taken from the vignette. The vignette specifies that all other aspects of the individual's life stay constant (including the time period), such that the only thing that can explain any difference in life satisfaction scores should be a difference in NHM visit frequency, provided that the respondents can answer the question accurately. We recognise that this may be a cognitively difficult task for the respondent but the previous vignettes literature does not provide any particular guidance or evidence on this issue.

In effect, the vignette is attempting to replicate an experimental setting whereby visit frequency is randomly assigned and all other factors are deemed to stay constant. An important difference is that in an actual experiment other (non-treatment) factors are balanced due to the act of randomisation whereas here we are relying on people to hold these other (non-treatment) factors constant, which is, of course, susceptible to bias.

In Rubin's (2005) terminology, the vignette means that each respondent has a value for life satisfaction under both potential states of the world.

This allows us to estimate wellbeing impact at the individual level rather than using expected values at the group level:

$$\tau_i \equiv Y_i(1) - Y_i(0) \quad (11)$$

Where  $\tau_i$  is the individual-level treatment effect;  $Y_i(1)$  is life satisfaction for individual with treatment (which here is the frequency of NHM visits in the vignette); and  $Y_i(0)$  is life satisfaction for individual  $i$  without treatment (which here is the actual level of NHM visits reported). We can assess the average causal effect for the sample as:

$$\sum_i \tau_i \quad (12)$$

which is identical to the average treatment effect (ATE) estimator from a standard experiment:

$$ATE \equiv \sum_i \tau_i \equiv E[Y_i(1)|D=1] - E[Y_i(0)|D=0] \quad (13)$$

where  $D$  takes on a value of 1 if the individual is treated.

$\sum_i \tau_i$  will be our estimate of  $\beta_2$  in equation (10).

The vignette study therefore offers another way of deriving values per visit based on life satisfaction impacts. There are relative advantages and disadvantages of the hybrid contingent-wellbeing valuation method and the vignette study. The former is simpler to administer and estimate and can be undertaken with small samples. The vignette-based WV requires larger samples for small activities and events and it requires some off-model analysis to convert impacts into monetary values using the framework set out in Annex C. However, it is arguably less likely to suffer from issues related to strategic bias and it does not ask individuals to assess the value of museum visits in monetary terms and so also reduces the likelihood of non-response or protest values in relation to the WTA question (a small number of respondents in our survey stated that a decrease in their life satisfaction could not be compensated in monetary or financial terms at all). Instead, in the vignette study respondents are required to think about how the cultural engagement impacts on their subjective wellbeing – arguably a less controversial question.

The values from the vignette study can be compared to the values derived from the hybrid contingent-wellbeing valuation and also to help assess the validity of the standard WV approach for cultural organisations. All three methods based on the WV approach have their strengths and weaknesses and should therefore be viewed as complements which help us explore in greater detail the mechanics and validity of WV. This is important for the development of the WV method for valuation in the cultural sector and also for our task of comparing CV and WV.

### 5.3. Implementation

We conducted **two surveys** with a total sample of 1,616 respondents. An **on-site exit survey** was administered at the Natural History Museum by a professional survey company (Morris, Hargreaves, McIntyre) in person, using an iPad, to 616 visitors at the Natural History Museum (14<sup>th</sup> October – 7<sup>th</sup> November 2014). The visitor survey was conducted both during half term and non-half term weeks. An **online survey** was implemented on a general population sample of 1,000 respondents using a web panel (5<sup>th</sup> November – 10<sup>th</sup> November 2014).

## 6. Case study one: Results

### 6.1. Socio-demographics

Table 6.1 summarises key socio-economic characteristics across the samples. For the online population survey, quotas were set on gender, age, and region to be in line with the target populations. For the case of the visitor survey we applied NHM visitor population weighting to the data.

**Table 6.1: Sample socio-economic characteristics**

	Visitor survey (weighted)	Visitor survey (unweighted)	Online survey
Male (%)	51	48	49
Age (mean)*	39	40	48
Household income (£, mean)*	39,000	39,000	30,000
Dependent children under 16 years (%)*	35	34	26
Married/with partner (%)	40	42	46
University education (%)*	62	61	35
In employment (full-time, part-time, self-employed) (%)*	72	69	54
Living in London (%)*	47	46	6
Health (good, very good, excellent)(%)*	94	93	71
Member (heritage, conservation, environmental org) (%)	17	18	22
Total	616	616	1,000

Notes: Gross annual household income; averages computed using the midpoints of the income and age categories. Legend: \*  $p < 0.05$  significant difference between surveys ( $t$ -test)

The proportion of women was just over half of the surveyed respondents in each survey (Table 6.1). The visitors were relatively younger, wealthier, more educated, much more likely to be in employment, more likely to have dependent children, and in better health. As expected, the proportion of those living in London was much higher amongst the sample of visitors. These differences were significant at the 5% level. Somewhat surprisingly, membership in heritage, conservation or environmental organisations was slightly higher amongst the general population.

### 6.2. Museum visits

Tables 6.2 and 6.3 summarise information about respondents' visits to museums. The results show that 65% of the general population claimed to have visited the NHM London at least once in their lifetime. 28% said that they had visited at least once within the last 12 months (of those who said they had visited NHM London in their lifetime, only  $n=18$  said they had not done so in the last year). As expected, awareness and visitation was much lower for the NHM site at Tring. The vast majority of respondents in both surveys had been taken to museums as children and had visited other museums and galleries in the last 12 months (Table 6.2).

**Table 6.2: Museum visits information**

	Visitor survey	Online survey
<b>Visited the NHM at least once in lifetime (%)</b>		
London	100	65
Tring	-	8
<b>Visited the NHM in the last 12 months (%)</b>		
London	100	28
Tring	-	7
<b>Enjoyment of visits to the NHM (4 or 5 on the scale 1-5, where 5 is enjoyed a lot) (%)</b>		
London*	94	89
Tring	-	87
<b>Familiarity with basic NHM information (very or extremely familiar) (%)</b>		
London*	19	33
Tring	-	11
Familiarity with information about NHM research and conservation work (very or extremely familiar) (%)	18	18
Visited other museums and galleries in last 12 months (%)*	85	63
Taken to museums and galleries as a child (%)	81	81
Total	616	1,000

Notes: Visitor survey figures weighted. Legend: \*  $p < 0.05$  significant difference between surveys (t-test)

The vast majority of those who had visited the NHM in London and Tring said they had highly enjoyed their visit (Table 6.2). But only a minority of respondents in each survey said they were very or extremely familiar with the basic information presented on the NHM in London. Only 11% of the general population sample was familiar with the information on the NHM in Tring. Just under 20% of respondents were familiar with the information on the NHM's research and conservation work in both surveys (Table 6.2).

Amongst the sample of visitors to the NHM in London, 41% were visiting with children. On average, visitors spent two hours on site, and as many as 43% said they had made a donation. The main reasons given for why visitors had come to the NHM London were: to enjoy the permanent collections (63%); to enjoy the museum's special temporary collections and exhibitions (23%), and to entertain and/or educate children (22%). The three most enjoyable areas of the NHM museum in London identified were the dinosaurs gallery (53%), the mammals and blue whale gallery (45%) and the volcanoes and earthquake gallery (31%) (Table 6.3). Differences in enjoyment between half term and non-half term visits were only statistically significant for the volcanoes and earthquake gallery and the restaurants and cafés, both of which were associated with higher enjoyment ratings during half term.

**Table 6.3: Most enjoyable areas of the museum (visitor survey)**

	%	Half term	Non-half term
Dinosaurs gallery	53	50	54
Mammals and blue whale gallery	45	44	45
Volcanoes and Earthquake gallery*	31	37	28
Minerals gallery (rocks, minerals, gems, crystals and meteorites)	24	25	24
Fossils galleries	19	22	17
Paid exhibitions	17	17	17
Darwin Centre	16	13	17
Birds gallery	14	14	14
Fishes, amphibians and reptiles gallery	14	12	15
The museum building and architecture (e.g. entrance hall)	10	7	11
Treasures in the Cadogan gallery	7	5	8
Restaurants or cafes*	5	2	7
Total	616	207	409

Legend: \*  $p < 0.05$  significant difference between surveys (*t*-test)

### 6.3. Attitudes

Visitors' general attitudes towards culture and the arts are depicted in Table 6.4. Some 18% of respondents in the general population survey and 25% in the visitor survey ranked heritage, arts, museums and culture amongst their three top priorities for public spending (from a list that included areas such as education, health, pensions, housing, environment, safety, transport, national defence and the economy). A large majority of respondents agreed or strongly agreed that the Natural History Museum was a national treasure that should be protected for future generations, and had a value even for those people who do not visit. Only around a quarter of the general population agreed or strongly agreed that there were more important things to spend money on than protecting UK museums such as the Natural History Museum. Interestingly, amongst the sample of visitors, 66% agree or strongly agreed with paying a small entrance fee to make sure the displays were kept up to date, clean and in working condition.

**Table 6.4: Attitudes towards culture and the NHM**

	Visitor survey	Online survey
Heritage, arts, museums and culture amongst the 3 top priorities for public spending (%)*	26	18
The Natural History Museum is a national treasure to be preserved for future generations (% agree or strongly agree)*	99	87
There are more important things to spend money on than protecting UK museums such as the Natural History Museum (% agree or strongly agree)	-	39
The Natural History Museum has a value even for those people who do not visit (% agree or strongly agree)*	84	69
I would not mind paying a small entrance fee to make sure the displays are kept up to date, clean and in working condition (% agree or strongly agree)	66	-

Legend: \*  $p < 0.05$  significant difference between surveys (*t*-test)

## 6.4. Contingent valuation analysis

### 6.4.1. Summary willingness to pay measures

Table 6.5 separates out those who are willing to pay something in principle. In the visitor survey, 85% of respondents said they would or maybe in principle would be willing to pay an entrance fee to access the NHM (which is broadly consistent with the generally favourable attitudes towards entrance fees depicted in Table 6.4); while in the online survey, 78% of respondents said they would (or would maybe) be willing to pay something (in the form of a donation) towards the continued existence of the NHM. When asked the actual WTP question, and faced with a list of possible amounts, a total of 22% of those visitors saying they were willing to pay for an individual ticket and 44% of those saying they were willing to pay for a family ticket were found to have a zero willingness to pay, while a total of 29% of the general population sample expressed a zero WTP (Table 6.6).

**Table 6.5. Willingness to pay**

	Visitor survey		Online survey	
	N.	%	N.	%
Yes	397	64.45	327	32.70
Maybe	130	20.94	452	45.20
No	90	14.61	221	22.10
Total	616	100	1,000	100

Table 6.6 shows descriptive statistics for each WTP measure, including mean and median WTP amount. In the visitor survey, mean willingness to pay for entry in the NHM was found to be £6.87 per individual entry ticket and, consistent with expectations, a higher £12.31 per family ticket (when controlling for group size, the WTP between those buying individual and family tickets was significantly different at  $p < 0.05$ ). When family ticket WTP was divided by the number of people in the group, and then added to the individual entrance fee WTP, we found that mean WTP for entry combined across the visitor survey was £6.65, very close to the mean WTP for individual tickets only. Median WTP was £6 for each type of ticket.

**Table 6.6. Mean and median WTP**

Survey	WTP variable	N	Low 95%	Mean (£)	High 95%	Median (£)	Max (£)	Zeros (%)
Visitor	Individual entrance fee	503	6.38	6.87*	7.36	6.00	32.50	21.7
Visitor	Family entrance fee	203	10.38	12.31*	14.25	6.00	52.50	44.3
Visitor	Combined WTP for access to the NHM (entrance fee)	616	6.23	6.65	7.06	6.00	32.50	17.7
Visitor	Visitor donation	616	2.59	2.78	2.98	2.25	17.50	14.3
Visitor	Visitor donation, if agreed to pay entrance fee	506	2.83	3.05*	3.27	2.25	17.50	10.1
Visitor	Visitor donation, if refused to pay entrance fee	109	1.18	1.53*	1.88	1.25	11.00	33.9
Online	Annual donation	996	7.21	8.29	9.37	4.25	175	29.0
Online	Annual donation (Visitors)	283	10.69	13.70*	16.72	5.50	175	14.1
Online	Annual donation (Non-visitors)	713	5.26	6.14*	7.02	2.75	112.50	34.9

Note: No responses in the binary choice WTP question ( $n=90$ ) were assigned in full to both the individual and family ticket variables, accounting for the difference in aggregate observations between them. Legend: \*  $p < 0.05$  significant difference between survey subsamples ( $t$ -test).

Visitor mean willingness to pay a donation on top of the entry ticket to support the research and conservation work of the NHM was found to be £2.78. As a proportion this is around 40% of the price people are willing to pay as an entrance fee.

Median WTP was £2.25. As expected, mean willingness to pay a donation was almost double for those who had already agreed to pay an entrance fee (£3.05) than for those who had refused one (£1.53). This difference in WTP was significant at the 5% level. Similarly, 34% of those who had refused to pay an entrance fee provided zero value donation responses, compared with only 10% of those who had agreed to pay an entrance fee (Table 6.6).

In the online survey, mean willingness to pay an annual donation to secure the future of the museum's collections, conservation work, and scientific research programme was £8.29. Median WTP was £4.25. As expected, mean WTP for those who had visited the NHM in London or Tring the last 12 months (£13.70) was higher than for those who had not visited (£6.14), significant at the 5% level (Table 6.6).

Distributions of willingness to pay values for each payment mechanism are presented in Annex D1.

**Table 6.7. Mean WTP by different socioeconomic groups (sample size in brackets)**

Socio-demographic characteristics	Visitor (use) WTP: Entrance fee	Visitor (non-use) WTP: Donation	General population mean WTP
Age: Under 50	£6.35* (465)	£2.75 (465)	£9.50* (560)
Age: Above 50	£7.56* (151)	£2.87 (151)	£6.74* (436)
Gender: Female	£6.54 (320)	£2.81 (320)	£7.82 (505)
Gender: Male	£6.76 (296)	£2.76 (296)	£8.77 (491)
Parental status: Dependent children	£6.23* (191)	£3.36* (191)	£8.75 (261)
Parental status: No dependent children	£7.43* (372)	£2.70* (371)	£8.13 (735)
Health status (self-reported): 'Good', 'Very good' or 'Excellent'	£6.64 (569)	£2.78 (569)	£8.92 (706)
Health status (self-reported): 'Fair' or 'Poor'	£6.42 (41)	£2.60 (41)	£6.76 (286)
Education: Degree and above	£6.92 (374)	£2.84 (374)	£10.64* (351)
Education: Up to degree level	£6.20 (238)	£2.67 (238)	£7.01* (645)
Income: <£33,400 per annum	£5.97* (251)	£2.43* (251)	£6.85* (596)
Income: >£33,400 per annum	£7.15* (349)	£3.04* (348)	£10.44* (400)
Employed (including self-employed and part-time)	£6.83 (420)	£2.93* (420)	£9.06 (538)
Unemployed	£5.54 (13)	£1.75 (13)	£6.09 (54)
Student	£5.56 (87)	£2.44 (87)	£14.44 (55)
Retired	£7.21 (71)	£2.46 (71)	£5.94 (224)
Inactive/unpaid family worker	£7.25 (2)	£0.33 (2)	£7.42 (125)
London	£6.09* (276)	£2.51* (275)	£14.48* (60)
East England	£6.80 (56)	£3.05 (56)	£6.78 (111)
East Midlands	£9.04 (27)	£4.18 (27)	£11.91 (76)
Northeast	£4.65 (13)	£1.08 (13)	£4.15 (36)
Northwest	£8.56 (24)	£2.90 (24)	£8.17 (127)
Southeast	£6.62 (118)	£2.64 (118)	£8.13 (147)
Southwest	£7.69 (34)	£3.57 (34)	£6.71 (115)
West Midlands	£8.18 (16)	£2.89 (16)	£11.20 (94)
Yorkshire & Humber	£9.27 (10)	£4.28 (10)	£5.82 (96)
Scotland	£7.90 (17)	£3.49 (17)	£5.05 (81)
Wales	£5.54 (3)	£8.42 (3)	£8.53 (52)
Mean WTP	£6.65 (616)	£2.78 (616)	£8.29 (996)

Note: some sociodemographic variables do not sum exactly to the full population sample due to missing responses for these questions in the survey.

Legend: \*  $p < 0.05$  significant difference between survey subsamples (t-test).

We find that mean WTP differed significantly with key demographic characteristics (Table 6.7). In particular, mean WTP for entry by visitors to the NHM was significantly higher among the over 50s, those with income above the population median, those with no dependent children, and those living in London. This is consistent with our prior expectation that those with most spending capability and lower budget constraints would generally be more willing to pay a higher entry fee. We found that general population WTP was significantly higher among those under 50, those with education at degree level and above, those in the top half income bracket, and those living in London.

In terms of geographical differences in WTP, with the exception of those the North East, visitors from the North and Midlands have a higher use value than Londoners and those from the East of England. This may be in part because a visit may be seen as more 'special' for those travelling longer distances compared with those that can visit at any time because of their proximity to the site (see below for further discussion). Non-use value, in contrast, is higher for Londoners. This confirms the theoretical expectation that non-use values are based around the institution itself, rather than distance travelled or any special event value.

#### **6.4.2 Validity tests**

In this section, we assess the validity of the WTP findings and provide supporting evidence that the WTP results can be interpreted as reflecting the welfare changes associated with visiting the NHM and with the range of services it provides. There are a number of ways to test the validity of the results. Here, we discuss content validity, theoretical validity and the bias reduction mechanisms included in the survey design (specifically to tackle the issue of hypothetical bias).

##### **6.4.2.1 Content validity**

Content validity tests look at the adequacy, realism and neutrality of the survey instrument as well as at respondents' understanding, perception and reactions to the questionnaire. Additionally, the rate of protest provides valuable information on how respondents reacted to the scenarios.

We conducted stakeholder meetings at the project inception stage with the NHM's Director and senior management team to agree the valuation scenarios for the contingent valuations (for both the visitor and online surveys). We then undertook extensive testing of the draft survey instruments. We used a face-to-face pilot survey followed by in-depth cognitive debriefing about all parts of the questionnaire with thirty visitors at the NHM; and an online pilot survey with some cognitive follow-ups on key parts of the questionnaire with 65 panel respondents, in both cases mimicking the conditions in which the final surveys would be implemented.

On the visitor pilot survey, in terms of difficulty, 91% of respondents found the survey to be 'ok' or 'easy', with none describing it as 'very hard'. Roughly 82% didn't think there were any questions that were too sensitive. Finally, the cognitive debriefing showed that only 27% of respondents found the survey length to be 'ok', while 55% found it 'a little long'. As a result the final questionnaire was substantively shortened and some of the longer questions dropped or edited down.

Pilot respondents were asked for their views on key and potentially complex parts of the questionnaire, such as the valuation and the wellbeing vignette questions. In terms of the valuation, some 88% of the sample found the scenario of government cuts to be 'realistic' or 'quite likely'. When asked about their certainty of paying the amount stated, 38% said they would definitely pay, 58% said they would most likely pay, and 4% were not sure. Respondents were also considering substitutes: when asked if an entrance fee at the NHM would reduce the number of other museums/galleries they could afford to visit, 57% said yes. In terms of the wellbeing vignette, 92% of the sample claimed to have understood it fully, and 88% did not find it to be complicated. However, 50% thought it was too long which led to some changes in this section of the questionnaire. An overwhelming 90% said they were very or quite certain about the subjective wellbeing scores they gave.

As regards the general population survey, the cognitive debriefing showed that 86% of respondents found the survey length to be 'ok'. Only 1.5% found it to be 'very long'. In terms of difficulty, 98.5% of respondents found the survey to be 'ok' or 'easy', with only 1.5% describing it as 'very hard'. Roughly 97% did not think that the survey used unclear or unusual language and 90% didn't think there were any questions that they felt uncomfortable answering. Moreover, 95% of the pilot sample agreed that the survey provided them with sufficient information.

In terms of the valuation section, 82% of respondents agreed that the survey provided enough information about the work of the Natural History Museum to answer the CV questions, and 86% found the NHM photos used in the survey instrument to be helpful for picturing its work and activities. Some 82% of the sample found the scenario of government cuts to be 'realistic' or 'quite likely'. Crucially, 80% of the sample did not find it hard to select a monetary value that they would be willing to pay as a donation to maintain the NHM and its work (and the remaining 20% found it only 'a little hard'). In addition, 74% stated that they fully understood the less common WTA scenario. In terms of the wellbeing vignette, 92% of the sample claimed to have understood it fully. Only 11% and 17% of respondents found the vignette question difficult or too long, respectively.

These findings indicate that our pilot survey instruments were generally well received and the public did not find them difficult to understand. Only minor changes were deemed to be required at this stage, mostly aimed at shortening the visitor survey.

Finally, in terms of protest answers, follow-up questions to the WTP scenarios show that only a very small proportion of respondents could be classified as protesters, i.e. stated a zero WTP when in fact they might value the change being proposed. In the visitor survey, only 10 respondents (less than 2% of the total) stated a zero WTP to the entrance fee question because they either didn't agree with charging for admission or they thought they paid enough for museums already through their taxes. In the general population survey, only 81 people (8% of the total) gave possible protest reasons for not offering a donation (e.g. because they paid enough already through other means such as taxes, because the NHM should charge entrance fees rather than ask for donations, or that the government should pay). Overall, these relatively small protest rates bode well for the validity of our estimated WTP measures.

#### **6.4.2.2 Theoretical validity**

The credibility of estimated values from CV studies is commonly assessed by examining their theoretical validity (Bateman et al. 2002). Theoretical validity examines whether the relationship between WTP and other indicators are in accordance with expectations. Some of these indicators are predictors from economic theory, while others reflect empirical regularities, which seem intuitively correct, from introspection and reasoned thought, and which have been found to hold across a large number of studies (Bateman et al. 2002). An example of the former is a positive relationship between WTP and income. An example of the latter is the effect on WTP of indicators such as use of museums or attitudes towards museums.

Testing for theoretical validity is typically done by estimating a bid function, exploring how WTP responds to respondent characteristics and other variables collected in the survey. In particular, we are interested in whether variables for which there are prior expectations are both significant in determining WTP, and affect values in the way that is expected. Put another way, if key variables are found to be either statistically insignificant or, most importantly, to affect WTP in unexpected and illogical ways, this casts doubt on the theoretical validity of results.

The set of variables included in the models is presented in Table 6.8. In our choice of predictors of WTP, we followed the recommendations of Bateman et al. (2002), which are common practice in modern applications of CV. In particular, we included a range of standard *socio-economic variables* (i.e. gender, age, children, education and income), variables relating to the *use of the good* (i.e. previous visits, time spent), *attitudinal variables* (i.e. enjoyment of the NHM, attitudes towards the NHM and public spending on culture), *distance-related variables* (distance from home to the NHM) and so on. For some variables there were no clear priors on the sign of the impact; that is, for example, in some of these cases the relationship with WTP could plausibly be positive or negative (e.g. having dependent children).

**Table 6.8. Determinants of WTP**

Variable	Interpretation	Mean
Gender (male)	Dummy variable (1-male; 0-female)	0.49
Age (log)	Log of age (mid-point of intervals)	3.72
Education (university)	Dummy variable (1-university degree or higher; 0-otherwise)	0.45
Income (log)	Log of annual household income before tax (mid-point of intervals)	3.25
Children	Dummy variable (1-dependent children under 16; 0-otherwise)	0.45
Distance travelled/distance to NHM (log)	Log of geographical distance from respondent postcode to NHM (km)	4.34
Individual ticket	Dummy variable (1-would pay for individual ticket; 0-would pay for family ticket)	0.79
Time spent in museum (log) (visitor survey only)	Time (hours) spent in NHM (log)	1.10
Half term (visitor survey only)	Dummy variable (1-visit during half term; 0-otherwise)	0.13
Weekend	Dummy variable (1-visit during weekend; 0-otherwise)	0.13
Weekend during half term (visitor survey only)	Dummy variable (1-visit during weekend of half term; 0-otherwise)	0.07
Visits to NHM in the last year	Continuous	0.87
NHM enjoyment (visitor survey only)	Likert scale 1-5 where 1 is 'did not enjoy at all' and 5 is 'enjoyed a lot'	4.47
Likely to visit NHM again (online survey only)	Likert scale 1-5 where 1 is 'not at likely', and 5 is 'very likely'	
Familiarity with NHM	Likert scale 1-5 where 1 is 'not at all' familiar, and 5 is 'very familiar'	2.76
Familiarity with NHM research and conservation work	Likert scale 1-5 where 1 is 'not at all' familiar, and 5 is 'very familiar'	2.34
Heard of NHM Tring (online survey only)	Dummy variable (1-heard of NHM at Tring; 0-not heard of NHM at Tring)	0.14
Recently attended other museums	Dummy variable (1-yes; 0-no)	0.71
Membership	Dummy variable (1-a member of a heritage, conservation, environmental or other organisation; 0-member of no organisation)	0.20
Closure of TL would reduce my life satisfaction	Dummy variable (1-closure of TL would reduce my life satisfaction; 0-closure of TL would have no effect on my satisfaction)	0.26
Cultural engagement: Art attendance growing up	Binary	0.81
Agreement: 'The NHM is a national treasure to be preserved for future generations'	Likert scale 1-5 where 1 is 'strongly disagree' and 5 is 'strongly agree'	4.49
Agreement: 'The NHM has a value to non-visitors'	Likert scale 1-5 where 1 is 'strongly disagree' and 5 is 'strongly agree'	4.01

Further descriptions of all of the conditioning variables used in the analysis can be found in Annex B.

**(i) Visitor WTP for access to the NHM (entrance fee)**

Table 6.9 shows the results for Models 1, 2 and 3 for museum visitors' willingness to pay an entrance fee (individual or family), controlling for a range of factors.

Income is, as expected, a significant and positive driver of higher willingness to pay across all three models (Bateman et al., 2002; Carson, 2012; Mourato et al., 2002).

Distance travelled turns out also to be a significant and positive driver. More qualitative work is needed to fully understand this result, but one plausible explanation is that a visit may be seen as more 'special' for those travelling longer distances compared with those that can visit at any time because of their proximity to the site. Time spent at the museum is not a significant driver of WTP.

Those who indicated in their responses that they would likely purchase an individual ticket have significantly higher WTP than those indicating they would purchase a family ticket. Plausibly, this may reflect greater budgetary constraints on those with families to support. Equally plausible, however, is that finding could be an artefact of the calculation method we used whereby we divided stated WTP by the size of the visiting group (in real-life situations, family tickets usually cover no more than two adults and two children).

Visiting the NHM at the weekend is a significant *negative* driver of willingness to pay. Possible interpretations include higher levels of congestion at the NHM at the weekend compared with midweek. Visiting during half term school holidays, or the interactive effect of weekends during half term, are not significant drivers of WTP. The lower willingness to pay may also reflect the time constraints of weekend tourists with busy sight-seeing schedules.

As expected, visit enjoyment is a significant positive driver of willingness to pay, suggesting that those who enjoy their time in the museum are, other things being equal, more willing to pay a hypothetical entrance fee to visit again.

In Model 3, the statement that "the Natural History Museum is a national treasure to be preserved for future generations" is a significant and positive driver of respondents' higher willingness to pay, capturing general positive attitudes towards the NHM. Variables included as more general proxies for cultural engagement are not significant drivers of willingness to pay an entrance fee.

**Table 6.9. Drivers of visitor willingness to pay for NHM access (individual and family entrance fee combined)**

Variable	Model 1 Coefficient	Model 2 Coefficient	Model 3 Coefficient
Gender (male)	0.799*	0.826*	0.903**
Age (log)	0.09	-0.48	-0.29
Education (university)	0.471	0.395	0.262
Income (log)	1.085***	1.141***	1.142***
Children	-0.066	-0.018	-0.055
Distance travelled (log)	0.444***	0.377**	0.380**
Individual ticket	3.823***	4.167***	4.025***
Time spent in museum (log)		0.7	0.492
Half term		-0.057	-0.076
Weekend		-1.273**	-1.224**
Weekend during half term		1.071	1.111
Visits to NHM in the last year		-0.102	-0.159
NHM enjoyment		0.697**	0.648**
Familiarity with NHM		0.257	0.152
Recently attended other museums		-0.05	0
Membership		0.562	0.43
Closure of NHM would reduce my life satisfaction			0.504
Cultural engagement: Art attendance growing up			-0.273
Agreement: 'The NHM is a national treasure to be preserved for future generations'			1.399***
Constant	-0.967	-3.393	-9.619***
Observations	469	468	468
r <sup>2</sup>	0.136	0.162	0.186

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at <10%. Reference group: (i) for gender ref = female; (ii) for education ref = all qualifications under Degree; (iii) for children ref = no children; (iv) for weekend ref = weekday; (v) for membership ref = membership of no organisation. Sample is restricted to England, Wales and Scotland and those aged 16 and over. Heteroskedasticity-robust standard errors.

### (ii) Visitor willingness to donate towards NHM's conservation and research work (top-up to entrance fee)

Table 6.10 shows the results for Models 1, 2 and 3 for museum visitors' willingness to pay a donation (as a top-up to the entrance fee) towards conservation and research work, controlling for a range of factors.

Income is again a significant and positive driver of willingness to pay in all three models.

In contrast with the previous model (Table 6.9), having dependent children is now a significant *positive* driver of willingness to pay. This suggests that the value individuals assign to NHM's research and conservation work has an element of bequest value (which would not be the case in their valuation of a visit). Distance travelled is again a significant positive determinant of WTP.

Time spent in the museum is in this case a strong positive driver of willingness to donate to support the NHM's conservation and research work, suggesting that the value people place on this aspect of the NHM is influenced by their recent experiences at the museum.

Familiarity with the research and conservation work of the NHM is statistically significant too in Model 2 but not in Model 3, suggesting that prior knowledge of the non-use work that NHM does is a weak driver of WTP.

Agreement with the statement that closure of the NHM for one year would reduce life satisfaction is a significant positive driver of WTP. This suggests that perceived personal effect of the NHM is an important factor in the value visitors place in the institution, which again conforms to our prior expectations.

Agreement with the statement that 'The NHM has a value to non-visitors' is a significant driver of willingness to donate, which adds support to the hypothesis that the values stated for donations are based around non-use attributes such as conservation and research.

**Table 6.10. Drivers of visitor willingness to donate for NHM's conservation and research work (top-up to entrance fee)**

Variable	Model 1 Coefficient	Model 2 Coefficient	Model 3 Coefficient
Gender (male)	0.089	0.046	0.049
Age (log)	-0.111	-0.4	-0.367
Education (university)	0.178	0.026	-0.019
Income (log)	0.356**	0.337**	0.350**
Children	0.485*	0.455*	0.491*
Distance travelled (log)	0.189**	0.162**	0.193***
Time spent in museum (log)		0.909**	0.850**
Visits to NHM in the last year		-0.031	-0.067
NHM enjoyment		0.006	-0.039
Familiarity with NHM research and conservation work		0.226**	0.112
Recently attended other museums		0.126	-0.029
Membership		0.436	0.356
Closure of NHM would reduce my life satisfaction			0.809**
Agreement: 'The NHM has a value to non-visitors'			0.490***
Constant	1.181	0.811	-0.906
Observations	544	541	539
r <sup>2</sup>	0.04	0.068	0.105

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at <10%. Reference group: (i) for gender ref = female; (ii) for education ref = all qualifications under Degree; (iii) for children ref = no children; (iv) for membership ref = membership of no organisation. Sample is restricted to England, Wales and Scotland and those aged 16 and over. Heteroskedasticity-robust standard errors.

### (iii) General population willingness to donate towards NHM conservation and research work (annual donation)

Table 6.11 shows the results for Models 1, 2 and 3 for the general population survey willingness to pay an annual donation towards conservation and research work of the NHM, controlling for a range of factors.

Income is a significant positive driver of willingness to pay a donation in all models. Higher education (degree or higher degree level) is also positive and statistically significant. This is consistent with previous contingent valuation studies of cultural

institutions which find that higher income and better educated individuals are more likely to pay more to support the work of cultural institutions (Noonan, 2003). Having dependent children is a significant negative determinant of WTP. Distance from the NHM is not a significant determinant, which reflects the non-use nature of many of the values being elicited.

Number of visits to the NHM in London in the last year is a significant positive driver of WTP, consistent with our expectation that those who use a cultural institution more often would value it more highly. Having heard of the NHM at Tring is a significant positive driver of WTP. This is consistent with the intuition that valuing a cultural institution like the NHM is contingent on one's knowledge of its existence and appreciation of its work.

As expected, given that the online survey also captures option values, the likelihood of visiting NHM in the future turns out also to be a significant positive driver of WTP.

Having attended museums and art galleries when growing up is a significant positive determinant of WTP too. This variable was designed as a proxy for an individual's cultural engagement. Having recently attended alternative museums is a significant negative determinant of WTP, consistent with findings from some previous studies that knowledge and experience of alternative sources of cultural value reduces WTP for the institution under study.

The statement that "The NHM has a value to non-visitors" remains a significant positive driver of willingness to donate, which accords with the non-use value elements that we have sought to capture in the online survey scenario.

**Table 6.11. Drivers of general population willingness to donate for NHM's conservation and research work (annual donation)**

Variable	Model 1 Coefficient	Model 2 Coefficient	Model 3 Coefficient
Gender (male)	0.737	-0.334	0.152
Age (log)	-3.120**	-0.544	-1.326
Education (university)	2.907**	2.222**	2.000*
Income (log)	2.097**	1.630**	1.461*
Children	-0.768	-3.305**	-2.731*
Distance to NHM (log)	-0.89	0.269	0.528
Visits to NHM (London) in the last year		3.408**	3.226**
Likely to visit NHM in future		1.413***	0.837*
Familiarity with NHM research and conservation work		0.443	0.228
Heard of NHM Tring		2.731	3.209*
Recently attended other museums		-1.423	-1.764*
Membership		0.677	1.092
Cultural engagement: Art attendance growing up			1.545*
Agreement: 'The NHM has a value to non-visitors'			2.385***
Constant	18.092**	-2.073	-9.015
Observations	995	982	949
r <sup>2</sup>	0.038	0.103	0.122

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at <10. Reference group: (i) for gender ref = female; (ii) for education ref = all qualifications under Degree; (iii) for children ref = no children; (iv) for weekend ref = weekday; (v) for membership ref = membership of no organisation. Controls are included for experimental branches of cheap talk/oath script in online survey. Sample is restricted to England, Wales and Scotland and those aged 16 and over. Heteroskedasticity-robust standard errors.

Taken as a whole, these findings evidently lend strong credence to the theoretical validity of our WTP estimates.

### 6.4.2.3 Hypothetical bias corrections

We included a number of bias reduction mechanisms in the survey design specifically to tackle the issue of hypothetical bias: in particular, certainty measures after the donation questions, and cheap talk and oath entreaties.

#### (i) Certainty

As discussed in Section 3.1, the effect of certainty on WTP has been found in some previous studies to be negative, suggesting that it is easier to be certain about paying small amounts (Bedate et al., 2009). However, this may not always be the case: it may equally be that those responding with a large amount have thought more deeply about the true value of the change being proposed, and that their higher value responses are therefore more considered and thoughtful than the lower value responses of others.

**Table 6.12. Impact of respondent certainty that they would pay the stated amount on willingness to pay**

Variable	Visitor survey: Use value Coefficient	General population survey Coefficient
Certain (very certain)	0.861*	1.605
Gender (male)	0.548	0.094
Age (log)	-1.073	-1.49
Education (university)	0.187	2.335*
Income (log)	1.409***	1.504
Children	-1.867***	-3.455*
Distance travelled (log)	0.300*	0.355
Time spent in museum (log)	-0.291	-
Half term	0.554	-
Weekend	-1.121*	-
Weekend during half term	0.141	-
Visits to NHM in the last year	0.025	3.275**
NHM enjoyment	0.573*	-
Likely to visit NHM in future	-	0.276
Familiarity with NHM	0.308	0.001
Heard of Tring	-	3.782*
Recently attended other museums	-0.504	-2.560**
Membership	0.274	0.416
Closure of NHM would reduce my life satisfaction	-0.227	-
Cultural engagement: Art attendance growing up	-0.889	1.312
Agreement: 'The NHM is a national treasure to be preserved for future generations'	1.481***	-
Agreement: 'The NHM has a value to non-visitors'	-	2.429***
Public spending on heritage, arts, museums and culture as priority	0.364	1.977
Constant	-3.837	-3.605
Observations	541	739
r <sup>2</sup>	0.134	0.109

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at <10%. Reference group: (i) for gender ref = female; (ii) for education ref = all qualifications under Degree; (iii) for children ref = no children; (iv) for weekend ref = weekday; (v) for membership ref = membership of no organisation. Sample is restricted to England, Wales and Scotland and those aged 16 and over. Heteroskedasticity-robust standard errors.

Table 6.12 shows that the level of certainty turns out in fact to be a significantly *positive* determinant of the stated donation in the NHM visitor survey, though not in the online survey. (Table 6.13 shows that certainty levels were reasonably high in both sets of survey responses).

**Table 6.13. Level of respondent certainty of their actual willingness to pay**

Level of certainty	Visitor survey %	General population survey %	Whole sample %
Not certain	11.83	16.17	14.26
Somewhat certain	40.52	60.33	51.58
Very certain	47.65	23.49	34.17
Observations	617	779	1,396

*Note: Respondents who answered No to the dichotomous WTP question in the general population survey were not asked the certainty of their response. This accounts for the lower number of total observations in this group.*

## (ii) Cheap talk and oath scripts entreaties

Table 6.14 shows that the oath script on its own had a significant negative effect on stated willingness to pay in the general population study. This suggests that the presence of an oath script reminding respondents to tell the truth succeeded in keeping elicited values lower. The cheap talk on its own also had a negative effect, though not statistically significant. Clearly, since the real value of donations is not observable, we cannot be sure that the value obtained with entreaties is any closer to true WTP than the value without entreaties. Nonetheless, the results are suggestive that the oath script entreaty may be a promising way to address hypothetical biases in our survey.

**Table 6.14. Entreaties experiments**

Variable	Coefficient
Oaths script and cheap talk	0.021
Cheap talk only	-0.107
Oath script only	-2.013*
Constant	-9.015
Observations	949
r <sup>2</sup>	0.122

*Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at <10%. Reference = no cheap talk or oath script. Sample is restricted to England, Wales and Scotland and those aged 16 and over. OLS regressions, controlled for a range of socio-demographic and experience factors (not shown). Heteroskedasticity-robust standard errors.*

## 6.5. Wellbeing analysis and valuation

### 6.5.1. Wellbeing summary statistics

**Table 6.15 Mean wellbeing scores across the visitor and general population surveys**

	Mean life satisfaction	Mean happiness (last hour)	Mean purpose (last hour)
Overall	7.035	7.096	7.079
Visitors	7.791	8.037	8.292
General population	6.569	6.513	6.332

All measures of wellbeing were on average higher in the visitor survey (Table 6.15).

**Table 6.16. Mean scores for happiness and purpose in the last hour by different activities at the Natural History Museum**

Activity	Mean happiness (last hour)	Mean purpose (last hour)
Paid exhibition/gallery	8.5	8.66
Darwin Centre	8.409	8.409
Free exhibition/gallery	8.018	8.263
Restaurants/cafes	7.759	8.321
Spending time in non-exhibition area	7.136	7.682
Other	5.6	5.6

The wellbeing ranking of NHM activities over the past hour in terms of average levels of happiness and purpose were the same across both measures, with attendance at paid exhibitions and the Darwin Centre being the highest.

### 6.5.2. Momentary wellbeing analysis

Tables 6.17 and 6.18 show the results for equation (8) for happiness and purpose in the last hour. Note that the number of observations used in the model may be lower than the full sample size of 1,616 due to non-response in some of the variables (this caveat applies to all of the models presented).

After controlling for a range of socio-demographic factors as well as current levels of life satisfaction, NHM activities rank highly in terms of associations with happiness. Being in the Darwin Centre or in a paid exhibition at the NHM has a higher positive association with happiness than any other activity. The highest ranking non-NHM activity is listening to music.

**Table 6.17 Association between activities and happiness in the last hour**

Natural History Museum	Coefficient	Natural History Museum	Coefficient
Regular exhibition	0.975***	Paid exhibition	1.702***
Darwin Centre	1.796***	Relaxing	0.336
Restaurant	0.743**	Other	-1.843
Other activities	Coefficient	Other activities	Coefficient
Internet	0.249	Care adults	0.121
Care children	0.623*	Computer games	1.237***
Cultural activities	0.847**	Eating	0.133
Hobbies	0.515	Housework	0.218
Reading/TV	0.820***	Sick in bed	-0.725
Sleeping	-0.730*	Sports	0.947***
Talking	1.014***	Travel	0.245
Cooking	0.846***	Music	1.284***
Something else	-0.027	Observations	1512
r <sup>2</sup>	0.565		

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at <10%. Activity refers to the main activity in the last hour. The following control variables included in the model but not shown here: gender, health, income, marital status, education, parental status, employment status, social networks, life satisfaction. Reference case for activities is working/studying. Heteroskedasticity-robust standard errors.

After controlling for a range of socio-demographic factors, as well as for current levels of life satisfaction, NHM activities also rank highly in terms of associations with purpose (Table 6.18). Being in the Darwin Centre or in a paid exhibition at the NHM has a higher positive association with purpose than any other activity. The highest ranking non-NHM activity is 'other cultural activities'.

**Table 6.18. Association between activities and purpose in the last hour**

Natural History Museum	Coefficient	Natural History Museum	Coefficient
Regular exhibition	1.331***	Paid exhibition	2.058***
Darwin Centre	1.805***	Relaxing	0.993**
Restaurant	1.411***	Other	-1.57
Other activities	Coefficient	Other activities	Coefficient
Internet	-0.01	Care adults	0.844
Care children	0.019	Computer games	1.368***
Cultural activities	1.398***	Eating	0.077
Hobbies	0.557	Housework	-0.007
Reading/TV	-0.327	Sick in bed	-0.917
Sleeping	-1.052**	Sports	1.275***
Talking	1.133***	Travel	-0.23
Cooking	0.784***	Music	1.249***
Something else	0.636**	Observations	1511
r <sup>2</sup>	0.51		

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at <10%. Activity refers to the main activity in the last hour. The following control variables included in the model but not shown here: gender, health, income, marital status, education, parental status, employment status, social networks, life satisfaction. Reference case for activities is working/studying. Heteroskedasticity-robust standard errors.

These results provide evidence that cultural activities are strongly associated with higher levels of wellbeing in the moment; people feel better and have a higher sense of purpose when engaged in cultural activities. Previous studies have shown positive associations between cultural activities and wellbeing at a more global level – in other words, between visits over a period of time (such as a year) and global measures of wellbeing such as life satisfaction. The results here can be seen as providing supporting evidence at a more micro level, looking at specific activities and momentary wellbeing, and are consistent with findings from Fujiwara and MacKerron (2015) who use an ESM dataset (Mappiness) and also find that individual cultural activities rank very highly, in terms of impacts on happiness and relaxation, in the moment.

### 6.5.3. Wellbeing valuation analysis

#### (i) Standard wellbeing valuation approach

Table 6.19 shows the results from estimating equation (9). This is the association between engagement with the NHM and life satisfaction after controlling for the main determinants of wellbeing. We look at the effect of being a visitor at the NHM over the past 12 months – i.e., 'Visited NHM' is a binary variable that equals 1 if the individual had visited the NHM at least once over the previous 12 months.

Visiting the NHM turns out to have a large positive association with life satisfaction after controlling for a wide range of socio-demographic variables. The estimated effect is, however, implausibly large: when valued using the WV procedure described in Annex C it produces a value of £6,003, almost equivalent to the impact of being in full-time employment (around £7,000)!<sup>10</sup> If

<sup>10</sup> The value for employment is broadly in line with other WV studies in this area (see Fujiwara, 2013).

we divide the coefficient by the average number of visits (0.87) it suggests that visiting the NHM twice in a year is equivalent to being in full-time employment in terms of impacts on life satisfaction. This is despite our being able to control for a wide range of variables to deal with self-selection into NHM engagement.

**Table 6.19. Association between visiting the NHM (any number of times) and life satisfaction**

Museums variables	Coefficient	Museums variables	Coefficient
Visited NHM	0.373**	Member NHM	0.092
Museums/galleries growing up	-0.069	Training/classes arts	0.123
Live near museum/gallery	0.036	Friends/family go museums/galleries	0.260**
Control variables	Coefficient	Control variables	Coefficient
Online survey	-0.514***	Male	0.111
Age (20-29)	-0.087	Age (30-39)	-0.265
Age (40-49)	-0.288	Age (50-59)	-0.063
Age (60-69)	0.286	Age (over 70)	0.642
Married	0.735***	Civil partner	0.361
Separated	-0.335	Divorced	0.165
Widowed	0.531*	GCSE	-0.024
A level	-0.232	Professional qualification	-0.37
Degree	-0.213	Higher degree	-0.284
Self-employed	0.239	Full-time	0.554**
Part-time	0.633**	Student	0.498*
Home care	0.403	Retiree	0.799**
Health excellent	2.100***	Health very good	1.751***
Health good	1.081***	Health fair	0.695**
Income (£15-19k)	0.262	Income (£20-29k)	0.520***
Income (£30-39k)	0.333*	Income (£40-49k)	0.644***
Income (£50-59k)	0.332	Income (£60-79k)	0.593***
Income (£80-99k)	0.405	Income (£100-149k)	1.281***
Income (£150k+)	1.428***	Children	-0.008
Socialises most days	0.612***	Socialises weekly	0.413***
Constant	4.189***	Observations	1241
r <sup>2</sup>	0.33		

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at 10%. Reference case is people who did not visit the NHM at all in the past 12 months. Heteroskedasticity-robust standard errors.

Tables 6.20 and 6.21 explore these results in greater detail. Table 6.20 looks at the association between visiting the NHM once in the past year and life satisfaction. The reference group is people who have not visited at all. Surprisingly, the coefficient on NHM visits is not significant in this case. In Table 6.21 we see that more frequent NHM visitors (more than one visit per year) do not have higher life satisfaction than people that visit the NHM just once per year either.

**Table 6.20. Association between visiting the NHM once in the past year and life satisfaction**

Museums variables	Coefficient	Museums variables	Coefficient
Visited NHM once	0.277	Member NHM	0.221
Museums/galleries growing up	-0.058	Training/classes arts	-0.006
Live near museum/gallery	0.148	Friends/family go museums/galleries	0.209*
Control variables	Coefficient	Control variables	Coefficient
Online survey	-0.716***	Male	0.072
Age (20-29)	-0.064	Age (30-39)	-0.275
Age (40-49)	-0.191	Age (50-59)	0.063
Age (60-69)	0.396	Age (over 70)	0.779
Married	0.723***	Civil partner	0.17
Separated	-0.19	Divorced	0.104
Widowed	0.519*	GCSE	-0.271
A level	-0.46	Professional qualification	-0.58
Degree	-0.454	Higher degree	-0.407
Self-employed	0.159	Full-time	0.477**
Part-time	0.598**	Student	0.572*
Home care	0.472	Retiree	0.694**
Health excellent	2.057***	Health very good	1.774***
Health good	1.117***	Health fair	0.691**
Income (£15-19k)	0.313	Income (£20-29k)	0.575***
Income (£30-39k)	0.268	Income (£40-49k)	0.704***
Income (£50-59k)	0.563**	Income (£60-79k)	0.573**
Income (£80-99k)	0.382	Income (£100-149k)	1.461***
Income (£150k+)	1.513***	Children	-0.024
Socialises most days	0.557***	Socialises weekly	0.493***
Constant	4.574***	Observations	1036
r <sup>2</sup>	0.328		

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at 10%. Reference case is people who did not visit the NHM at all in the past 12 months. Heteroskedasticity-robust standard errors.

**Table 6.21. Association between NHM visit frequency and life satisfaction in comparison to visiting the NHM once**

Museums variables	Coefficient	Museums variables	Coefficient
Visited NHM 2-4 times	0.044	Visited NHM 5-8 times	-0.257
Visited NHM 9-12 times	-0.144	Visited NHM 17-20 times	0.061
Member NHM	0.037	Museums/galleries growing up	-0.259
Training/classes arts	0.292**	Live near museum/gallery	-0.044
Friends/family go museums/galleries	0.203		
Control variables	Coefficient	Control variables	Coefficient
Online survey	-0.558***	Male	-0.237**
Age (20-29)	-0.075	Age (30-39)	-0.291
Age (40-49)	-0.424	Age (50-59)	-0.082
Age (60-69)	0.431	Age (over 70)	0.209
Married	0.379**	Civil partner	0.146
Separated	-0.935*	Divorced	-0.156
Widowed	0.062	GCSE	0.68
A level	0.085	Professional qualification	0.031
Degree	0.128	Higher degree	0.054
Self-employed	0.779**	Full-time	0.941***
Part-time	1.097***	Student	0.753*
Home care	-0.031	Retiree	1.052**
Health excellent	0.56	Health very good	0.244
Health good	-0.392	Health fair	-0.56
Income (£15-19k)	0.014	Income (£20-29k)	0.281
Income (£30-39k)	0.126	Income (£40-49k)	0.351
Income (£50-59k)	0.016	Income (£60-79k)	0.459*
Income (£80-99k)	0.306	Income (£100-149k)	1.012***
Income (£150k+)	1.247***	Children	0.293*
Socialises most days	0.563***	Socialises weekly	0.07
Constant	6.345***	Observations	703
r <sup>2</sup>	0.254		

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at 10%. Reference case is people who visited the NHM once in the past 12 months. People with no visits dropped from the analysis. Heteroskedasticity-robust standard errors.

These results suggest that it is not accurate to interpret the significant large coefficient on NHM visits in Table 6.19 as impacts due to visits, since that coefficient is not sensitive to visit frequency. It may conceivably indicate selection bias. However, considering that we have controlled for a rich variety of selection factors that typically are not available when using secondary

datasets in studies of museums and wellbeing, and, importantly, that the coefficients for the other variables in the model are in line with the empirical wellbeing literature, we might speculate that these biases should be less severe in this study compared with others. In that case the large coefficient on NHM visits arguably requires further explanation.

One possibility is if priming effects have distorted our survey responses. Taken at face value, the results suggest there is a substantive difference in life satisfaction between someone that visits the NHM (even if just once) and someone that doesn't visit at all. Although we have controlled for more readily observed differences (e.g., education and income), other differences are hidden. If, say, those who visit do so because they generally have a higher level of interest or affection for the NHM (including, possibly, higher option and non-use values), then being reminded about the NHM during the survey may boost their reported life satisfaction disproportionately, in a similar way to some of the other contextual factors discussed in Section 2.1.2, such as the weather and a football team winning before the survey.

This would offer one speculative reason why the coefficient on NHM visits is not responsive to visit frequency, and yet is substantially (and implausibly) larger than the findings of previous studies of the wellbeing impact of museum visits.

To explore this hypothesis further, we undertook further analysis (not presented here) where we estimated the effect of NHM visit frequency on life satisfaction for people that had visited one or more times using the same set of control variables as the models in Tables 6.19, 6.20 and 6.21. In effect, we replaced the Visited NHM variable in Table 6.19 with a visit frequency variable and restricted the sample to those who had visited at least once. This model removes the source of the priming effect we describe above because there are no people with zero visits in the sample (the priming effect impacts on people who visit the NHM *compared to people that don't visit at all*). The coefficient on NHM visit frequency in this model turns out to be small, at about 0.01, and statistically insignificant. This supports our interpretation that the large positive coefficient on Visited NHM in Table 6.19 is not in fact capturing the true effect of NHM visitation on life satisfaction.

The possibility of such priming effects is, of course, just one of many alternative explanations offered by behavioural economics, but they do fit the pattern of the results. Whatever the explanation, the conclusion is clear: implementing life satisfaction WV in the case of the NHM was not successful (and we note the possibility that priming effects may be the source of the problem).

We now turn to the vignette and the hybrid contingent-wellbeing valuation results to assess whether they provide additional methods for valuing cultural activities, using the fundamental theoretical basis of the WV approach, and which could complement CV when applied in a cultural institution.

### (ii) Hybrid contingent-wellbeing valuation

Table 6.22 summarises the results from presenting the hybrid contingent-wellbeing valuation scenario to the survey respondents. In the visitor survey, 17% of respondents said that closure of the NHM would have an effect on life satisfaction, compared with 32% in the general population survey. 83% of respondents to the visitor survey said that closure of the NHM would have very little effect on life satisfaction. This figure was lower (67%) in the general population survey. Less than 1% of respondents to the general population survey said that the closure would increase their life satisfaction, while no respondents indicated that closure would increase their life satisfaction in the visitor survey. We treated responses that closure would increase life satisfaction from our analysis as protest votes and removed them from the analysis.

**Table 6.22 Willingness to Accept closure of the NHM**

	Visitor survey		General population		Total	
	N	%	N	%	N	%
The closure would have very little effect on my life satisfaction	503	83.14	440	67.18	943	74.84
The closure would reduce my life satisfaction	102	16.86	209	31.91	311	24.68
The closure would increase my life satisfaction	0	0	6	0.92	6	0.48
Total	605	100	655	100	1,274	100

*Note: The Willingness to Accept closure of the NHM question in the General Population survey was in fact only asked to those who lived in the London area or had visited the NHM at least once in their lifetime.*

Table 6.23 shows the average WTA values for closure of the NHM for one year to keep life satisfaction constant.

This question was in fact asked only of the 25% of respondents that said their life satisfaction would decrease due to the closure of the NHM. For this group, the mean level of WTA for one-year closure turned out to be £27.02. For the 75% of respondents for whom the closure would have a very small impact on life satisfaction, we assume that WTA closure of the NHM for one year was a nominal £1 for this group (which equates to about the bottom third of WTA values for the group that said their life satisfaction would decrease).

**For the full sample** (i.e., excluding those who indicated that the closure would increase their life satisfaction), **the mean level of WTA for closure of the NHM for a year is £7.45.**<sup>11</sup> If we assume that the individual would continue to visit the NHM at the same rate as they had reported for the previous year then it is possible to calculate **a value per NHM visit which is £6.89 per visit** (based on an average annual visit frequency of 1.08 for those answering the WTA question). This lies slightly above the individual and family ticket WTP from the CV (at £6.87 and £12.31 respectively). In this case, the use of the explicit reference to keeping life satisfaction constant seems to have eradicated the well-documented positive WTA-WTP disparity that we discussed in Section 5.2.2.2.

The results also allow us to look at the life satisfaction impact of additional visits to the NHM, which we speculate was not possible with the standard WV method due to priming effects. In the WTA scenario, people are explicitly made to think about the use value of visits, as they are told that they are not able to visit the museum at all for one year.

Whereas in the standard WV analysis visit frequency had no estimated effect on life satisfaction, in the WTA scenario people seemingly place a realistic value on an additional NHM visit. As the task is to compensate for a loss in life satisfaction due to the museum’s closure this implies that people actually predict a small impact on their life satisfaction from NHM visits.

It is possible, through a process of ‘reverse engineering’, to back out from the WTA value what the predicted life satisfaction impact would be. In other words, we can replicate a coefficient for NHM visits based on the WTA values. This is achieved by using equation (10’):

$$\text{Value of NHM engagement} = \beta_2 / \beta_1 \quad (10')$$

where we know the left hand side of the equation for a single visit is £6.89 and the impact of income ( $\beta_1$ ) from the income model in Annex C ( $\beta_1 = 1.1$ ). The value for  $\beta_2$  which solves this equation turns out to be 0.0023.<sup>12</sup> In other words, the WTA value together with the estimate for  $\beta_1$  from the income model suggest that people predict that closure of the NHM for one year would reduce life satisfaction by 0.0023 points. Given an average visit frequency of 1.08 this translates in to a life satisfaction impact of 0.0026 points per visit.

**Table 6.23. WTA values based on life satisfaction**

Mean WTA	Average visits	WTA (per visit)	Estimated LS coefficient (annual)	Estimated LS coefficient (visit)
£7.45	1.08	£6.89	0.0023	0.0026

**(iii) Vignette-based wellbeing valuation**

The vignette study can also be used to derive estimates of the impact of NHM visits on life satisfaction and the value of these visits. Table 6.24 shows the average treatment effect (ATE) for life satisfaction, estimated using equation (12), and the NHM visit frequency associated with the ATE.

**Table 6.24 Increases in life satisfaction due to vignette visits**

Mean difference in LS ( $\sum_i \tau_i$ )	Mean difference in visits	Change in LS per visit	Vignette value per visit
0.0118	4.278	0.0028	£40.42

<sup>11</sup> Note that using a conservative estimate where those who indicate that the closure would have no effect on their satisfaction are recorded as having a £0 WTA, we find a mean WTA of £6.67 for the full sample.

<sup>12</sup> Note that we use the non-linear version of equation (10’) from Annex C (i.e., equation (C3)).

The vignette scenario involves, on average, about 4.3 more visits per year, and the vignette life satisfaction scores are, as expected, higher than the actual life satisfaction scores. Under the assumption, discussed in Section 5.2.2.2, that people are able to predict impacts on their life satisfaction, the difference in life satisfaction scores can be attributed to this difference in visits since no other factor changes across the two settings.

An increase in NHM visits – of around 4 visits – turns out to be associated with an increase in life satisfaction of 0.0118 points. **This is equivalent to an increase of 0.0028 life satisfaction points per visit, which has a value of about £40 per visit estimated using the WV approach** set out in the Annex C. This is estimated where  $\beta_2 = 0.0028$  and  $\beta_1 = 1.1$  (The result is not statistically significant at the 5% level, but arguably this reflects the small sample size).

The vignette study – like the hybrid contingent-wellbeing valuation – removes the possibility of a priming effect as respondents focus only on changes in visit frequency. However, the value and impact size per visit estimated from the vignette are larger (by a magnitude of about three) than the figures estimated from the WTA results.

The hybrid CV-WV method and vignette-based wellbeing valuation offer two new ways of estimating values associated with cultural organisations using the theory underlying WV. These findings at least suggest that the estimated values are higher under the vignette approach.

These results are broadly consistent with our study of momentary wellbeing: since cultural experiences, such as visiting the NHM, have large impacts on momentary wellbeing, we would expect there to be a positive, albeit small, effect of visit frequency on global measures of wellbeing as well. And this is indeed what we have found in both the hybrid CV-WV and vignette-based WV studies.

## 7. Case study one: Discussion and interpretation of the results

Table 7.1 summarises the main results from our case study at the Natural History Museum. Care must be taken when using and interpreting the range of values estimated. We have estimated three core types of WTP with different meanings depending on the context they refer to. Visitor use value can be divided between individuals who indicated they would pay for individual or family entrance. We provide a combined entrance fee WTP that includes *all visitors* surveyed. This value should be used as the primary estimate of direct use value for current visitors.

Visitor non-use value represents visitors' WTP an additional (one-off) donation to support the research and conservation work of the NHM. This is a supplementary WTP that must be interpreted alongside visitor use value. We note the conceptual and practical difficulties of separating direct use and non-use value among visitors. We cannot, therefore, exclude the possibility that visitor use values capture some element of non-use value, and vice versa.

The online survey captures non-use and option values for the general UK population (excluding Northern Ireland) as an annual donation. This value is conceptually distinct to both use and non-use values elicited in the visitor survey, referring to the range of values that the general public attach to a large international cultural institution. It is not, for example, possible to equate non-use value amongst visitors to non-use value in the general population, since the former is elicited via a one-off visitor donation, while the latter is elicited via an annual donation.

The values elicited in Table 7.1 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, and among visitors whether we are interested in direct use or non-use values, as well as the influence that the payment mechanism (entry fee or donation) has on the size of the value elicited.

**Table 7.1. NHM valuation results**

Survey	Use/non-use value	Valuation variable	Mean	Median
Visitor	Use	Individual entrance fee (WTP)	£6.87	£6.00
Visitor	Use	Family entrance fee (WTP)	£12.31	£6.00
	Use	Combined WTP for access to the NHM (entrance fee)	£6.65	£6.00
Visitor	Non-use (visitors)	Visitor donation (WTP)	£2.78	£2.25
Online	Non-use and option (general population)	Annual donation (WTP)	£8.29	£4.25
Visitor and online		Compensation for closure (per visit) (WTA)	£6.89	N/A
Visitor and online		Vignette wellbeing value (per visit)	£40.42	N/A

Note: values calculated using mid-point of intervals from the payment cards. Standard WV values are excluded due to the discussed problems concerning priming effects in the survey.

The study produced **visitor WTP values for access to the NHM of a plausible magnitude and comparable with prices currently charged for paid exhibitions in UK museums** (for example, for the NHM's Wildlife Photographer of the Year exhibition and for Coral Reefs: Secret Cities of the Sea, which opened in March 2015, these were £12.60 and £10 respectively). The WTP for a family ticket was about double the WTP for an individual ticket and **WTP amounts were found to be affected by theoretically relevant factors such as income, visitor enjoyment and positive attitudes towards the museum.**

**Mean levels of visitor donations (as a top-up to the entrance fee) for the NHM's conservation and research work were also of a realistic level** and driven by theoretically relevant factors. Specifically, **income and time spent in the museum** are significant positive drivers of higher hypothetical donations. The latter suggests that the value people place on the (mostly non-use) conservation and research work of the NHM is partly driven by recent direct use experiences. This suggests that despite the presence of separate information and familiarity questions on the non-use activities of the NHM, it was not possible to fully prevent the conflation of use and non-use value in the responses of museum visitors. As noted earlier, it is cognitively very difficult to disentangle different components of total economic value for a single individual. Moreover, it could well be the case that for some people a visit raises awareness of the importance of the conservation and research work carried out by the NHM in a way that is not captured by other control variables.

**Mean and median annual willingness to donate by the general population to support the work of the NHM were also of plausible magnitudes.** The results again accord strongly with theoretically relevant factors. **Income, likelihood of a future visit and familiarity with the NHM in Tring** were all predictors of higher willingness to pay. The latter is consistent with the idea that valuing a cultural institution like the NHM is contingent on one's knowledge of its existence and awareness of the value of its work. **Indicators of cultural engagement**, such as having attended art galleries and museums while growing up, were also significant positive drivers of WTP, suggesting that those who are most culturally engaged value the work of the NHM most highly. We found that **higher education was also a significant and positive driver of value**, consistent with previous contingent valuation studies on cultural institutions (Noonan, 2003).

**Agreement with statements about the non-use value of the NHM** was found to be a significant positive driver of willingness to pay in both the online and visitor surveys, where the scenario focused on the conservation and research work of the NHM, consistent with the fact that the donation scenarios were intended to capture elements of non-use values.

And crucially, **income was a positive and significant determinant of WTP** in all models, and all scenarios, as predicted by economic theory.

Overall, there were seemingly **few protest** responses. There was some indication of hypothetical bias in the general population survey that used an annual donation payment mechanism, because the **oath script** – which aims to reduce these types of biases – had a significant negative effect on WTP. In this case, within the constraints of our small sample size, we did not in contrast find a significant impact from cheap talk. The results point to the oath script as the more promising tool for addressing hypothetical bias in cultural engagement settings.

In order to help reduce informational constraints we used **best-practice methods** which included:

- Use of photos and detailed descriptions of the good/service plus payment vehicle.
- Checking familiarity with the various pieces of information provided during test surveys.
- A wide payment range in the payment ladder (informed by the results of survey testing) to reduce anchoring and range effects.

**An important consideration, and qualification, for our findings is that the stated WTP for entrance fees and donations is likely to be driven or primed to some extent by current paid exhibition prices and donation suggestions (which are usually around £3-£5 in UK museums).** And in this sense the plausibility of the results from the WTP questions is arguably not so surprising after all, as the approach intends to mimic what would happen if we had a real market for museum entry. An unanswered question remains as to whether the market prices anchor the stated WTP upwards or downwards, however. Against this, we note that the CV method may face the polar opposite problem in some cultural settings – that is, when there is no relevant market price that could act as a feasible anchor as in the case, say, of unique and iconic collections. This is an important area for future research.

**In sum, our study follows best practice in CV methodology and finds that the method is a viable approach for measuring use and to some extent non-use (and option) value in the context of a major national museum, the Natural History Museum. The method produces realistic values that vary in ways predicted by theory and by previous findings, across a range of scenarios, payment vehicles and population groups.**

Our analysis of the association between activities and happiness and purpose at the momentary level shows that **partaking in cultural activities and events at the NHM is also associated with significantly higher levels of wellbeing than other activities.** These results are in line with previous findings using ESM data that show that cultural activities rank very highly in terms of impacts on SWB. This provides some **rationale for why people are willing to pay to visit the NHM and also suggests there should be an association between NHM engagement and SWB over the longer term too.**

The analysis of the life satisfaction data shows that **visiting the NHM has an implausibly large impact on life satisfaction, however.** The magnitude of this effect is not supported by other studies of cultural activities and wellbeing, and is nearly as large as some major life events such as employment. Selection bias may always be an issue in life satisfaction models, as we have discussed, but we have controlled for a much richer array of selection factors in this analysis than previous studies of cultural activities and wellbeing – which typically use secondary data – have been able to do. **We speculate that the estimated large life satisfaction impact is driven by an artefact of the survey, namely a priming effect** whereby people who care for or like the NHM are being reminded of the importance of the NHM in the survey, which artificially inflates their wellbeing at the time of the survey (along the same lines as previous studies that have reported impacts of trivial contextual factors on life satisfaction).

The study provides evidence therefore that **life satisfaction wellbeing valuation is not suitable in instances where priming effects may be a problem.** In contrast, results for engagement with cultural activities using general national survey data (where this priming effect should not be an issue) have generated plausible wellbeing valuations (e.g., see Fujiwara et al., 2014). Based on the results of this study and previous WV studies, we might therefore tentatively conclude that **the WV method is more suitable for valuing engagement with culture over time and more generally than for discrete experiences with individual institutions.**

What we have called the **hybrid contingent-wellbeing valuation approach, i.e. the WTA scenario based on life satisfaction, for a scenario involving a hypothetical closure of the NHM for one year, gives plausible values per visit.** The hybrid design employed – where compensation should only *just* compensate respondents for an impact on wellbeing and crucially only for those that said their wellbeing would be negatively affected in the first place – seemingly has the added advantage of eradicating disparities between WTA and WTP valuations that are often found in the literature.

**The findings from applying the vignette-based WV may tentatively suggest a new method for valuing visits to cultural institutions.** The study design seemingly eradicated the priming effect that can affect standard WV analysis, and life satisfaction was responsive to visit frequency in the way expected. The estimated value per visit was around six times higher than the WTP and WTA equivalent values, however. As this is the first time that a vignette study design has been used in valuing culture – and notwithstanding that the approach worked less well in the case of the Tate Liverpool study (as we shall see) – our results suggest there is merit in further research.

In sum, our study of the NHM provides evidence that WTP and hybrid versions of the WTA question appear to perform well for valuing cultural engagement. Although WTP has now become the preferred monetary elicitation method in CV practice, there are times when WTA could be warranted. This may be when respondents come from very poor backgrounds, say, such that their WTP amounts are severely constrained and they feel uncomfortable about being asked to pay (even if they might be prepared to pay a small amount), and hence offer a protest zero. Another scenario which may warrant use of a WTA question is when property rights are such that respondents can be judged to have some *intrinsic* right to the good/service – and what's more they recognise this. This may be especially relevant for cultural activities and institutions. **In instances when a WTA payment structure is preferred our case study of the NHM shows that a hybrid approach grounded in the theory of WV can deliver plausible values in the context of cultural engagement.**

A further possible criticism of WV concerns the **limitation of life satisfaction as a measure of the broad concept of quality of life**. If preferences in fact represent a broader measure of wellbeing, then this should result in *smaller* values from WV relative to stated preference valuation. However, the results from the WTA and vignette studies – which represent the most comparable study design to the WTP question in this study – **do not support this claim**: values elicited based on the theory of WV and life satisfaction are of the same order of magnitude to the WTP values based on preferences, and if anything are a little higher.

**The conventional approach to WV is best suited to the valuation of non-marginal changes in states or conditions or for measuring the value of frequent activities where priming effects are not an issue.** These are precisely instances to which CV is not so well suited. However, cultural institutions seeking to carry out surveys on wellbeing with visitors and members should note the potential for problems like priming effects which can distort the impacts that can be assessed using global wellbeing measures. We have found that CV comes into its own in such cases. Our study with the NHM suggests that vignette-based WV has the potential to contribute in such cases too, but much more research is required.

## 8. Case study two: Tate Liverpool

Tate Liverpool is one of the largest galleries of modern and contemporary art in the UK outside of London. It is housed in the award-winning conversion of the nineteenth century Albert Dock on Liverpool's waterfront. The gallery displays modern and contemporary art from the Tate Collection. Tate Liverpool also has a family space in the Clore Learning Centre which is a hub of holiday and weekend activities, including Future Tate, a monthly art club for eleven- to fourteen-year olds.

In 2013, Tate Liverpool had nearly 600,000 visitors, and 41% of visitors to the gallery cited it as their reason for coming to the city. Tate Liverpool has connections with universities and art colleges in the city, and does a large amount of work within the local community, including partnerships with local organisations using art as a therapeutic activity for long-stay health patients and adults with special needs. Other programmes include its work with children's centres in deprived areas of Liverpool and within the galleries through Tate Liverpool's early-year specialists.

Entrance to the gallery and its displays of the national collection of modern and contemporary art is free. Admission is charged for entry to special temporary exhibitions. Tate Liverpool is part of the wider Tate Gallery; around 60 per cent of Tate Gallery's funding comes from non-government sources.<sup>13</sup>

Another publicly supported UK cultural institution, producing a range of market and non-market benefits, Tate Liverpool provides an interesting context for studying the applicability of economic valuation techniques, and is the subject of our second case study.

<sup>13</sup> [www.tate.org.uk/about/who-we-are/funding](http://www.tate.org.uk/about/who-we-are/funding)

## 9. Case study two: Survey methodology

Mirroring the procedure in the Natural History Museum case study, we conducted two surveys, each combining contingent valuation with wellbeing measurement. The first survey was a face-to-face exit survey of visitors to Tate Liverpool (N=526). The second was an online survey of the UK population excluding Northern Ireland (N=1,000). Visitor and general population (online) surveys were screened to exclude UK non-residents and those under sixteen years old. For the online survey, quotas for gender, age, and region were set using national UK averages from the Office for National Statistics' Annual Population Survey. On-site interviewers followed a protocol of randomly selecting one in every three people exiting the gallery. We then applied TL visitor weights for gender and age (assembled for the 2013-14 financial year and collected by random sampling) to the visitor survey data to correct for slight over-representation of older people and a slight gender imbalance (see weighting table, Annex G6).

The survey was again divided into four sections. The first section contained background questions on the size of respondents' visitor groups, number of children, use and enjoyment of Tate Liverpool, and whether they had made a donation during their visit. The visitor survey included specific questions on respondents' recent experiences in Tate Liverpool, which parts of the gallery they had most enjoyed, and their reasons for visiting. The online survey asked whether respondents had heard of, visited, and enjoyed Tate Liverpool and their likelihood of visiting again. Following issues identified in the NHM survey, which suggested some element of overstatement of recent visits to the cultural institution under study, we implemented calibrating questions in the online survey. For instance, we did not reveal that the focus of the study was in fact Tate Liverpool until after the frequency of visit questions had been asked. We also asked follow up questions to those who indicated that they had visited in the last 12 months, including the date of their last visit (in 3 month ranges), which included dates outside of the 12 month period, which were subsequently excluded from analysis.

As in the NHM study, both visitor and general population surveys asked respondents identical questions about membership of heritage, conservation or environmental organisations, and their use of other galleries and cultural sites in the last twelve months. They were also asked the same prioritization question on the allocation of government funding, and the same set of questions about their cultural engagement while growing up. A set of statements about the value of Tate Liverpool were presented on a five-point Likert scale (1, strongly disagree; 5, strongly agree).

In the second section, both sets of respondents were asked the following three SWB questions:

1. Overall, how satisfied are you with your life nowadays? ('Life satisfaction')
2. How happy were you feeling in the last hour or so? ('Happiness')
3. How worthwhile did what you were doing in the last hour or so feel? ('Purpose')

People responded on a scale of 0-10 where 0 = 'Not at all' and 10 = 'Completely'. In the visitor survey, respondents were asked which gallery activity they had been doing in the previous hour. In the online survey, we asked what activities the respondent had been doing in the previous hour, developed from the Mappiness iPhone Experience Sampling Method (ESM) (MacKerron and Mourato, 2013).

In both surveys we asked respondents if they thought they would be happier if they could visit Tate Liverpool more often. A life satisfaction vignette scenario was presented to all respondents. The vignette asked respondents to:

"Take your life as it currently is, but now imagine that you have visited Tate Liverpool **[once every [frequency] months (in other words, about [frequency] over the year)]** alone or with your family/friends. Assume that you usually spent one or two hours there per visit and vary the activities each time". Overall how satisfied would you be with your life?"

We randomised the visit frequency between once every [six/four/three/two/one month (s)], or twice a month in order to derive unbiased estimates of the impact of TL visits on life satisfaction.

The third section of the surveys presented respondents with identical information on Tate Liverpool, in terms of the services and visitor attractions it offered, and the work it undertakes outside the gallery in the wider community. Respondents were asked their familiarity with the information before this survey on a five-point Likert scale (1, not at all familiar; 5, extremely familiar).

Willingness to pay scenarios, payment mechanism, and elicitation method differed between the visitor and online survey. Below we describe the key components of the design of the CV and the SWB components of the surveys and their methods of analysis.

The fourth section asked a set of standard socio-demographic questions for use in the CV and SWB analysis.

## 9.1 Contingent valuation

The visitor survey contingent valuation scenarios were designed to uncover the following values:

- (A) the value of accessing Tate Liverpool (via a one-off donation).
- (B) the value of the work Tate Liverpool does outside of the gallery in the wider community (via a one-off donation, over and above any previous donation).

The general population survey contingent valuation scenarios were designed to uncover:

- (C) the value of Tate Liverpool, in terms of the work it does both inside and the gallery (via an annual donation).

The 'hybrid' contingent-wellbeing valuation question was asked of both visitors and the general population, designed to uncover:

- (D) the value of avoiding the closure of Tate Liverpool for one year (via a one-off cash compensation).

In terms of the Total Economic Value framework, (A) is a direct use value; (B) is a non-use value with possible elements of option value (in that respondents may benefit from the community outreach work of Tate Liverpool – we assume, however, that respondents have not directly used community outreach services in the past); (C) is similar to (B), i.e. a non-use value with possible elements of option value; (D) is likely to be mostly a use-value and an option value, as closure would prevent access and future access but not the on-going work of Tate Liverpool outside the gallery. The following section provides further details about these scenarios.

### 9.1.1 Contingent scenarios design

As in the NHM case study, the on-site visitor survey explored use values related to Tate Liverpool using stated preference methods. Respondents were asked to state the maximum they would be willing to pay as a donation to Tate Liverpool and how much compensation they would require if Tate Liverpool were to close for a period of time (and hence they could not visit). The compensation question was also designed to be used in the wellbeing analysis.

In the visitor survey we asked two contingent valuation questions.

The first, relating to use values, asked questions about the value of a visit to Tate Liverpool (see Box B1). We provided two paragraphs of information on the experiences that Tate Liverpool offers, including recreation, education, inspiration, and asked respondents about their familiarity with the information.

The payment vehicle used was a one-off donation, asked only of those over 16 years of age. We laid out a hypothetical scenario where galleries in the UK were considering alternative ways of funding its activities due to restrictions in government funding. In CV surveys in order to elicit welfare-consistent values it is vital that the hypothetical scenario be believable, and that it justifies introduction of the hypothetical payment mechanism (in this case a visitor donation). We presented respondents with a payment card offering them a range of monetary amounts they could choose from. This provided respondents with a visual aid and helped remove starting point bias (Bateman et al., 2002; Maddison and Foster, 2003; Maddison and Mourato, 2001).

**Box B1. Visitor survey donation (use) willingness to pay question**

In a hypothetical scenario, if there were restrictions in Government funding because of the current financial situation, Tate Liverpool might have to consider alternative ways of funding both:

- the maintenance and improvement of the presentation, display and facilities at the gallery in order to enhance visitor experience; and
- its community outreach programmes.

In the next set of questions, we are interested to know how much, if anything, you would be prepared to donate, on top of existing membership or temporary exhibition fees, to support each of these programmes (improvement of visitor experience and outreach activities) separately. It might be that you value both programmes, or just one of the programmes, or none of the programmes. I will ask you about willingness to donate to each programme in turn.

Firstly, would you in principle be prepared to make a one-off donation, even if only a small amount, to support the work that Tate Liverpool does *inside the gallery* to improve visitor experience? The donation would be collected via the existing donation boxes in the gallery, and would be used exclusively for maintenance and improvement of the presentation, display and facilities at the gallery.

Note that this donation would be *additional* to any membership or paid exhibition fees and to any other donation you may decide to make for other Tate Liverpool programmes.

IF YES OR MAYBE: In a moment we will ask the maximum amount you would be prepared to pay in principle, as a one-off donation, even if only a small amount, to support the work that Tate Liverpool does inside the gallery to improve the visitor experience.

Studies have shown that many people, answering surveys such as this one, say they are willing to pay more in donations than they would actually pay in reality. So please think about this question as if it were a real decision and you were actually making a donation for real. Please do not agree to pay an amount if you think you cannot afford it, if you feel you have paid enough already, or have other things to spend your money on (including support for other galleries and museums).

Also, in answering this question, please focus solely on how much *visiting* Tate Liverpool is worth to you. Please do not consider the value to you of other activities that Tate Liverpool may carry out outside the gallery, nor the value of the work carried out by Tate in London or Cornwall, nor the value of other galleries and museums. In this question, we are just interested in how much you feel you benefit from the displays, presentation and facilities at Tate Liverpool.

WTP values were elicited using a payment card with values ranging from £0 to £50.

As discussed earlier, the second contingent valuation question was intended to explore mostly non-use values, asking respondents to value the work performed by Tate Liverpool outside of the gallery. We provided respondents with a paragraph of information on the work Tate Liverpool does within the local community, including partnerships with local organisations using art as a therapeutic activity for long-stay health patients and adults with special needs, as well as work with children's centres in deprived areas of Liverpool and within the galleries through Tate Liverpool's early-year specialists. The payment vehicle was a one-off donation over and above any previous donation they had already indicated they would pay (see Box B2).

**Box B2. Visitor survey donation (non-use) willingness to pay question**

Second, if there were restrictions in Government funding because of the current financial situation, Tate Liverpool might consider alternative ways of funding its community outreach programmes.

Would you in principle be prepared to make a separate one-off donation, even if only a small amount, to support *the work that Tate Liverpool does with the wider community, using art to improve health and education outcomes in Liverpool?* The donation would be collected via the existing donation boxes in the gallery and would be used exclusively for Tate Liverpool's community outreach projects.

Note that this donation would be *over and above* what you already said you would donate (if anything) to improve the displays, presentation and facilities at the gallery, and overall visitor experience. It would be a separate donation for a different purpose. It would also be additional to any membership or paid exhibition fees.

Would you be prepared to make a one-off donation even if only a small amount, to support Tate Liverpool's continued community outreach work?

IF YES OR MAYBE: What is the maximum amount you would be prepared to pay as a *one off donation* even if only a small amount, to support Tate Liverpool's community outreach work?

In answering this question, please focus solely on how much you think Tate Liverpool's community outreach work is worth. Please do not agree to pay an amount if you think you cannot afford it, if you feel you have paid enough already, or have other things to spend your money on, and other ways to fund community development. Also please be realistic – consider your household budget and remember that there may be other things you would like to spend your money on.

WTP values were elicited using a payment card with values ranging from £0 to £50.

The online general population survey was intended to explore mostly non-use and option values. We asked respondents to imagine that Tate Liverpool was considering alternative ways of funding the work it does making modern and contemporary art accessible to the public, as well as supporting its work outside of the gallery (including its education programs, community programs, and its wider contribution to the city) (see Box B3). We provided respondents with information about the work of Tate Liverpool inside and outside the gallery, using illustrative photographs, and elicited how familiar they had been with the information beforehand. We used the payment vehicle of an annual donation.

**Box B3. Online survey donation (non-use) willingness to pay question**

In a hypothetical scenario, if there were restrictions in Government funding because of the current financial situation, Tate Liverpool might have to consider alternative ways of funding the work it does making modern and contemporary art accessible to the public, as well as supporting its work outside of the gallery (including its education programmes, community programmes, and its wider contribution to the city).

If Government funding ceases to be provided, Tate Liverpool would need to raise enough money in donations to support the work it does inside and outside the gallery.

For the next set of questions, please imagine a situation where Government funding for art galleries and museums has been cut. In the unlikely event that Government funding ceased, please think about how much Tate Liverpool would be worth to you and your family, if anything.

Would you in principle be prepared to make an *annual donation*, even if only a small amount, to help maintain and improve the gallery and to support the work that Tate Liverpool makes to the wider community?

IF YES OR MAYBE: Please think about how much Tate Liverpool is worth to you, if anything. What is the maximum you would be willing to pay, as a *donation*, per year, to secure the work that Tate Liverpool does inside the gallery, as well as to support the contributions that Tate Liverpool makes to the wider community?

Studies have shown that many people answering surveys such as this one, say they are willing to pay more in donations than they would actually pay in reality. So please think about this question as if it were a real decision and you were actually making a donation for real.

[RANDOMISE THIS PARAGRAPH TO SUBSET OF SAMPLE FOR CHEAP TALK/OATH SCRIPT EXPERIMENT:

Please do not agree to pay an amount if you think you cannot afford it, if you feel you have paid enough already, or have other things to spend your money on, and other ways to fund the arts and community work Tate Liverpool does. Also, the question is just about the work of Tate Liverpool and not about other galleries.]

WTP values were elicited using a payment card with values ranging from £0 to £150.

In both surveys, echoing what was done with the NHM, respondents were asked a certainty question on the donation amount that they had stated. Respondents were also asked to give reasons for their willingness, or not, to pay.

In designing the contingent valuation scenarios (for both the visitor and online surveys) we implemented the same best practice to attempt to deal with the known biases in CV, related to question wording and ordering, payment format, payment vehicle, prompts and use of images (Bateman et al., 2002). Respondents were 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 (Champ and Bishop, 2001; Cummings and Taylor, 1999). Certainty questions were asked after the two contingent valuation questions (Bedate et al., 2009; Champ and Bishop, 2001), and we collected follow-up information on the reasons that they were or were not willing to pay the amount stated (Bateman et al., 2002).

We ran a randomised experiment within the online survey to test best practice in reducing hypothetical bias in contingent valuation methods. We set up the same four-branch experiment as in the NHM case study for this purpose: a cheap talk script, an oath script, cheap talk and oath scripts in combination, and neither cheap talk or oath script.

### 9.1.2 Contingent valuation analysis

As described above, the visitor and general population surveys provided three separate WTP and one WTA payment mechanisms.

- A one-off donation for the work TL does inside the gallery asked to the full sample of n=526 visitors
- A one-off donation for the work TL does outside the gallery, on top of any previous donation, asked of the full sample of n=526 visitors
- An annual donation towards the work TL does inside and outside the gallery asked in the online general population survey (n=1,000)
- A cash compensation to all respondents across both surveys (n=1,526). The analysis of WTA is covered in detail in Section 5.2.2.2.

As we discussed previously, the 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). We calculated non-parametric mean and median WTP and WTA from the mid-point between the amount chosen on the card and the next amount up. Open-end responses produced no outliers above £500 to be removed as protest votes and all unanswered open-space responses were coded as missing.

The same regression model as in the NHM case study was used as the base for the WTP analyses:

$$WTP_i = \alpha + \beta_1 X_i + \varepsilon^i \quad (5)$$

where  $WTP_i$  is the amount the individual  $i$  has stated they are willing to pay,  $\alpha$  is the deterministic factor and  $\varepsilon$  is the error term containing unobserved factors that determine willingness to pay. In  $X_i$  we control for the observed determinants of willingness to pay (Bateman et al., 2002), including those that are theoretically expected to affect WTP as well as other factors that are known from the literature to have an effect.

We estimated three variants of equation (5) for each of the three WTP measures, meaning nine models in total.

- Model 1: Equation (5) with socio-demographic variables
- Model 2: Model 1 with variables capturing experience at the gallery
- Model 3: Model 2 with attitudes, opinions, and proxy variables for cultural engagement

We addressed heteroskedasticity using robust standard errors and tested for normality in the distribution of residuals using kernel density estimates (Annex G7).

We applied the same tests as in the NHM case study for the validity of our results. For instance, we tested for differences in WTP between those who indicated that they were very certain with the answers they gave, and those that were either uncertain or somewhat certain:

$$WTP_i = \alpha + \beta_1 Cert_i + \beta_2 X_i + \varepsilon_i \quad (6)$$

where  $WTP_i$  is the amount the individual  $i$  has stated they are willing to pay,  $Cert_i$  is the individual's stated certainty to pay that value,  $X_i$  controls for determinants of WTP in Model 1,  $\alpha$  is the deterministic factor and  $\varepsilon$  is the error term containing unobserved factors that influence WTP.

We explored possible protest bids by analysing the reasons given by respondents for being willing or not willing to pay a donation (Annex G3). Since the estimated number of protests was small (see Section 10.4.2.1), all responses were included in the analysis.

Following the procedure we adopted in the case of the NHM case study we tested for differences in stated WTP between individuals randomly shown one of the four experimental cheap-talk/oath branches in the online survey using the following model:

$$WTP_i = \alpha + \beta_1 CT_i + \beta_2 OA_i + \beta_3 CTOA_i + \beta_2 X_i + \varepsilon_i \quad (7)$$

where  $WTP_i$  was the amount the individual  $i$  had stated they were willing to pay,  $CT_i$  is a dummy variable that denotes whether an individual was presented with only the cheap talk script,  $OA_i$  is a dummy variable marking whether an individual was presented with an oath script only,  $CTOA_i$  captures when an individual was presented with the cheap talk and oath script together and,  $X_i$  controls for determinants of WTP in Model 1,  $\alpha$  is the deterministic factor and  $\varepsilon$  is the error term containing unobserved determinants.

This allowed us to quantify the additional effect of an oath instrument or cheap talk script on stated WTP, and to test the interactive effect of the oath and cheap talk together – hence guiding us on how to minimise hypothetical bias.

## 9.2 Subjective wellbeing

### 9.2.1 Subjective wellbeing design

Again replicating the NHM case study, both surveys used in the TL case study, asked respondents questions about their SWB and a set of socio-demographic variables. Variation in TL visit frequency across the full sample was used to assess the impact of TL visits and engagement on life satisfaction after controlling for a wide range of factors.

We again collected data on the individual's main activity over the past hour to assess the drivers of momentary wellbeing, as captured by their happiness and purpose.

A vignette was used to further assess the association between TL visits and life satisfaction. Specifically, respondents were asked to imagine a situation where they were able to visit TL more frequently and were asked what their level of life satisfaction would be, holding all other factors in their lives constant.

Results for associations with life satisfaction were, as before, used to derive monetary values using the wellbeing valuation approach. The questions on happiness and purpose in the moment provided supporting information for interpreting the results of the life satisfaction and contingent valuation analysis, and gave us a fine-grained take on the relationship between specific gallery activities and SWB.

The visitor and online surveys used in the Tate Liverpool case study can be found in Annex E of the paper.

### 9.2.2. Subjective wellbeing analysis

There are two aims of the SWB analysis in the TL case study:

- (i) To assess the impacts of TL activities and visits on momentary wellbeing.
- (ii) To estimate the values that people place on engagement with Tate Liverpool, using impact on (non-monetary) life satisfaction and wellbeing valuation to compare these values with values estimated using CV.

#### 9.2.2.1. Momentary wellbeing

We assess the relationship between different cultural and non-cultural activities and momentary SWB, measured as happiness and purpose, over a short period of time (the last hour) using the same econometric model we used in the NHM case study:

$$SWB_i = \alpha + \beta_1 A_i + \beta_2 X_i + \varepsilon_i \quad (8)$$

where  $SWB_i$  is individual  $i$ 's subjective wellbeing (happiness or purpose) over the previous hour;  $A_i$  is a list of activities that the individual undertook during this period (where working/studying is used as the reference case in the models);  $X_i$  is a vector of control variables and  $\varepsilon_i$  is an error term.

All the SWB models were estimated using ordinary least squares (OLS) regression analysis and heteroskedasticity-robust standard errors for the reasons discussed in Section 5.2.2.1.

#### 9.2.2.2. Wellbeing valuation

This section covers the same material as in section 5.2.2.2 in the NHM case study, but for ease of reference many of the salient points are repeated here.

##### (i) Standard wellbeing valuation approach

The standard WV approach uses the outputs from a life satisfaction model of the following type:

$$LS_i = \alpha + \beta_1 M_i + \beta_2 TL_i + \beta_3 X_i + \varepsilon_i \quad (9)$$

where  $LS_i$  is life satisfaction for individual  $i$ ;  $M_i$  is income;  $TL_i$  is a variable proxying engagement with Tate Liverpool; and  $X_i$  is a vector of other determinants of life satisfaction. In the analysis, TL engagement is represented by a variable denoting whether the individual has visited Tate Liverpool over the past year. As in the momentary wellbeing models,  $X_i$  include standard determinants of wellbeing (see Fujiwara and Campbell, 2011).

In addition, we control for whether the survey was taken online or in person to account for possible survey mode effects, and any unobserved differences present when the surveys were completed, such as in the weather. We also control for a set of factors that make selection into gallery engagement more likely. These are variables that have been found to be important determinants of engagement in galleries and cultural activities in previous studies: whether or not an individual is a member of Tate Liverpool;

whether they were taken to museums (or arts galleries) as a child; whether they took any classes or training in the arts; whether they live near a museum or gallery, and whether friends and family often go to museums or galleries. All these variables might also plausibly correlate with subjective wellbeing and therefore their exclusion from the model risks giving rise to omitted variable bias.

The economic value of TL engagement can be assessed by using the estimates of the coefficients for income and TL (respectively  $\beta_1$  and  $\beta_2$ ). The ratio of these two figures is an estimate of the value (CS) of TL engagement because it shows how much extra income is required to have the same impact on life satisfaction as TL engagement. Put another way, it is the amount of money required to keep individuals just as satisfied with life in the absence of visiting Tate Liverpool (i.e., to keep their wellbeing constant).

$$\text{Value of TL engagement} = \beta_2 / \beta_1 \quad (10)$$

This is a heuristic description of the approach we follow. In practice we use a more involved method for estimating  $\beta_1$ , which adopts a non-linear structure and which controls for selection bias, though the general principle of comparing the ratio of the two coefficients still applies. (See Annex C for the technical background and details of the statistical methodology employed in the WV analysis).

The analysis, as with most studies in this area, is necessarily based on observational datasets (i.e., where people have not been assigned to different conditions in a controlled experimental setting). Thus cause and effect relationships are *approximated* using statistical methods, as causation cannot be directly inferred and we cannot hope to capture all confounding factors which correlate with both TL engagement and wellbeing. Notwithstanding this difficulty, we follow best practice in WV by controlling for what we can in the regression analysis.

## (ii) Hybrid contingent-wellbeing valuation

In this approach, we ask respondents directly how much monetary compensation they would require if they were not able to visit Tate Liverpool for a period of time (one year) such that their life satisfaction would remain unaffected. Note that this is essentially a contingent valuation study using a WTA elicitation format. Since the approach makes explicit mention of 'wellbeing' (or life satisfaction in this case) however, we describe it as a hybrid contingent-wellbeing valuation approach.

In the TL case study we use the following question (Box B4):

### Box B4: Hybrid contingent-wellbeing valuation question

*For this next question, please imagine that Tate Liverpool had to **close to the public for one year** for vital maintenance work. No one would be able to visit any parts of the gallery during this period. All **other public galleries would remain open as usual**. Now don't worry, there are no plans for Tate Liverpool to close! But we would like you to think about what your life would be like if it did close for one year. How would the closure affect you level of life satisfaction?*

- The closure would have no effect on my life satisfaction*
- The closure would reduce my life satisfaction*
- The closure would increase my life satisfaction*

*If the respondent selects option 2 then,*

*You stated that the closure would reduce your life satisfaction. Your current level of life satisfaction is [xx – INSERT FIGURE FROM B1]. Assuming that nothing else in your life would change, think about your level of life satisfaction during the period of the closure?*

*Now imagine the following situation. Suppose that in order to compensate you for not being able to visit Tate Liverpool during one year, you were given a cash compensation. How much money would you have to receive, as a one-off payment, to give you the same **life satisfaction** that you have now (**not better nor worse, but just the same**) during this period until the gallery re-opened? Think about this for a moment please.*

WTA values were, as before, elicited using a payment card with values ranging from £0 to £150. We asked a set of follow up questions to establish the reasons why the respondent was/was not willing to accept monetary compensation for the closure of Tate Liverpool.

This follows the theoretical approach underlying WV, but the difference is that equation (10) is, in effect, internalised by the respondent. That is, instead of estimating the elements of equation (10) ( $\beta_1$  and  $\beta_2$ ) through data on people's actual visit

frequency and life satisfaction we pass on that task to the individual who needs to think about how her life satisfaction is impacted on by a restriction on TL visits (akin to  $\beta_2$ ) and how much money would compensate (offset) this effect (akin to  $\beta_1$ ).

Willingness to accept (WTA) compensation questions are sometimes used in CV, but normally respondents are not asked for compensation in terms of wellbeing/life satisfaction impacts – they are typically just asked for compensation *per se*. Our project is the first to assess whether constraining the WTA scenario by setting it explicitly in the context of changes in life satisfaction produces reasonable WTA values relative to WTP. The hypothesised mechanism for this is that respondents are directed explicitly to think about the WTA question within the framework of economic theory (i.e., compensating a welfare change). We note that our definition of welfare here (i.e., life satisfaction) differs from the standard definition of preference utility in economics, but we would still hypothesize an effect due simply to the explicit reminder to keep welfare constant.

This element of the study is not free of statistical problems, however. Although we do not estimate  $\beta_1$  and  $\beta_2$  statistically, biases will arise if people are unable to predict the impact of museum closure on life satisfaction. Indeed, a number of experiments have shown that people are unable to do this with accuracy (Kahneman, 2000). Kahneman and Snell (1992), for example, report that people find it very hard to predict how much pleasure they will derive from consuming even everyday goods such as music, yoghurt and ice cream. Gilbert (2007) and Kahneman and Snell (1992) attribute these types of findings to a 'presentism heuristic': people project current tastes and desires on to their predicted future tastes and preferences. Clearly though, preference-based valuation methods such as CV will also suffer from these prediction biases.

**(iii) Vignette-based wellbeing valuation**

In our third approach to WV, we use a vignette to overcome two possible problems with standard WV. The first problem is that if repeat visits to Tate Liverpool are not numerous there may not be enough variation across the sample to detect an impact of visits on life satisfaction and this creates problems for WV. The second problem is that it is difficult to establish causality with any confidence absent an experimental study.

A vignette may, at least in principle, address both problems by presenting individuals with a hypothetical life or event and asking them questions about this case. Anchoring vignettes are brief descriptions of hypothetical situations measuring a single concept, with responses rated on a Likert-type scale (King et al., 2004). Respondents may be questioned on each concept with both self-assessment items and hypothetical items that represent the range of possible levels of the variable being measured.

As in the NHM case study, our vignette study uses an anchoring vignette design to determine the effect of engaging with Tate Liverpool on life satisfaction, by presenting survey respondents with a varying number of visits to the gallery by a hypothetical individual and then asking what they think that individual's life satisfaction would be.

Specifically, we adopt a *first person vignette* (FPV), which takes the hypothetical element of a regular vignette and adds to it the current situation of the individual/respondent. It asks about the life satisfaction score the respondent would attribute to her life under some hypothetical scenario, where some element of her life has changed (hypothetically). The vignette question we use is,

*'Take your life as it currently is, but now imagine that you have visited Tate Liverpool [once every [frequency] months (in other words, about [frequency] over the past year)] alone or with your family/friends. Assume that you usually spent one or two hours there per visit and vary the activities each time'. Overall how satisfied would you be with your life?'*

Not at all										Completely
0	1	2	3	4	5	6	7	8	9	10

We vary the visit frequency randomly across the sample with a range of [once every SIX months]/[once every FOUR months]/[once every THREE months]/[once every TWO months]/[ONCE every month]/[TWICE every month].

We hypothesise that given that people are being asked essentially about their *own* lives, albeit with a minor change in circumstances, they should find it reasonably straightforward to respond to the SWB question as in third party reports, and certainly more easily than for an unfamiliar subject. Consistent with this, results from our pre-testing of the survey with visitors and an online population suggest that respondents found the vignette question generally easy to answer. Of the 73 test respondents 93% said that they had no problem with the clarity of the language in the vignette.

By adopting the FPV research design we aim to reduce (i) informational problems related to not knowing all aspects of the vignette subject's life, and (ii) focusing illusion bias, since respondents only focus on one individual life – their own (although focusing illusion cannot be totally eradicated since respondents naturally compare the life described in the vignette with their own lives) where the difference in TL visit frequency would get highlighted.

As in the hybrid contingent-wellbeing valuation, the vignette study provides an alternative approach to valuation using the WV method as a base. But this time, the estimation of the effect of TL visits ( $\beta_2$ ) is internalised by the respondent; we pass on the task to the individual who needs to think about how her life satisfaction is impacted on by changes in TL visit frequency. This still leaves the impact of income on life satisfaction ( $\beta_1$ ) to be estimated, however, and we do this separately using the WV methods set out in Annex C. That is, as discussed earlier, the vignette approach risks the same problems as does the hybrid contingent-wellbeing valuation, as people are asked to predict impacts on life satisfaction.

The vignette study is used to derive estimates of the impact of TL visits on life satisfaction and, given the estimated impact of income on life satisfaction, the value of these visits. Specifically, each survey respondent has (i) an actual level of life satisfaction associated with an actual level of TL visits, and (ii) a hypothetical level of life satisfaction associated with a hypothetical level of TL visits taken from the vignette. The vignette specifies that all other aspects of the individual's life stay constant (including the time period), such that the only thing that can explain any difference in life satisfaction scores should be a difference in TL visit frequency, provided that the respondents can answer the question accurately. We recognise that this may be a cognitively difficult task for the respondent but the previous vignettes literature does not provide any particular guidance or evidence on this issue.

In effect, the vignette is attempting to replicate an experimental setting whereby visit frequency is randomly assigned and all other factors are deemed to stay constant. An important difference is that in an actual experiment other (non-treatment) factors are balanced due to the act of randomisation whereas here we are relying on people to hold these other (non-treatment) factors constant, which of course is susceptible to bias.

In Rubin's (2005) terminology, the vignette means that each respondent has a value for life satisfaction under both potential states of the world.

This allows us to estimate wellbeing impact at the individual level rather than using expected values at the group level:

$$\tau_i \equiv Y_i(1) - Y_i(0) \quad (11)$$

Where  $\tau_i$  is the individual-level treatment effect; is life  $Y_i(1)$  satisfaction for individual  $i$  with treatment (which here is the frequency of TL visits in the vignette); and  $Y_i(0)$  is life satisfaction for individual  $i$  without treatment (which here is the actual level of TL visits reported). We can assess the average causal effect for the sample as:

$$\sum_i \tau_i \quad (12)$$

which is identical to the average treatment effect (ATE) estimator from a standard experiment:

$$ATE \equiv \sum_i \tau_i \equiv E[Y_i(1)|D=1] - E[Y_i(0)|D=0] \quad (13)$$

where  $D$  takes on a value of 1 if the individual is treated.

$\sum_i \tau_i$  will be our estimate of  $\beta_2$  in equation (10).

The vignette study therefore offers another way of deriving values per visit based on life satisfaction impacts.

All three methods based on the WV approach have their strengths and weaknesses and should therefore be viewed as complements which help us explore in greater detail the mechanics and validity of WV. This is important for the development of the WV method for valuation in the cultural sector and also for our task of comparing CV and WV.

### 9.3. Implementation

We conducted **two surveys** with a total sample of 1,526 respondents. An **on-site exit survey** was administered by a professional survey company (Morris, Hargreaves, McIntyre) in person, using an iPad, to 526 visitors at Tate Liverpool (2<sup>nd</sup> February – 8<sup>th</sup> April 2015). An **online survey** was implemented on a UK excluding N. Ireland general population sample of 1,000 respondents using a web panel (24<sup>th</sup>– 28<sup>th</sup> March 2015).

## 10. Case study two: Results

### 10.1. Socio-demographics

Table 10.1 summarises key socio-economic characteristics across the samples. For the general population survey, quotas were set on gender, age, and region to be in line with the target populations. For the case of the visitor survey we applied TL visitor population weighting to the data.

**Table 10.1: Sample socio-economic characteristics**

	Visitor survey (weighted)	Visitor survey (unweighted)	General population survey
Male (%)*	43	40	49
Age (mean)	47	47	45
Household income (£, mean)*	41,000	41,000	30,000
Dependent children under 16 years (%)*	15	14	30
Married/with partner (%)	43	43	46
University education (%)*	68	67	35
In employment (full-time, part-time, self-employed) (%)*	65	64	52
Living in Northwest (%)*	62	62	50
Health (good, very good, excellent)(%)*	92	92	74
Member (heritage, conservation, environmental org) (%)*	37	34	18
Total	526	526	1,000

*Notes: Gross annual household income; averages computed using the midpoints of the income and age categories. Legend: \*  $p < 0.05$  significant difference between surveys (t-test)*

The proportion of women was just over half of the surveyed respondents in each survey (Table 10.1). The visitors were relatively older, wealthier, more educated, more likely to be in employment, less likely to have dependent children, and in better health. As expected, the proportion of those living in the Northwest of England and members of heritage, conservation, or environmental organisation was much higher amongst the sample of visitors.

### 10.2. Gallery visits

Tables 10.2 and 10.3 summarise information about respondents' visits to museums. The results show that 25% of the general population had claimed to have visited Tate Liverpool at least once in their lifetime (this might appear implausibly high – see below). 10% said they had visited at least once within the last 12 months. 89% of Tate Liverpool visitors reported having visited other museums or galleries in the past 12 months compared to only 54% of the general population. The vast majority of respondents in both surveys were taken to museums as children (Table 10.2).

**Table 10.2: Museum visits information**

	Visitor survey			General population survey		
	Northwest	Rest of UK	Whole survey	Northwest	Rest of UK	Whole survey
Visited TL at least once in lifetime (%)	-	-	-	36	35	25
Visited TL in the last 12 months (%)	-	-	-	14	6	10
Enjoyment of visits to TL (4 or 5 on the scale 1-5, where 5 is enjoyed a lot) (%)	81	75	79	80	84	81
Familiarity with basic TL information (very or extremely familiar) (%)*	34	15	27	16	15	15
Familiarity with information about TL community outreach work (very or extremely familiar) (%)	13	4	10	9	15	12
Visited other museums and galleries in last 12 months (%)*	92	84	89	52	56	54
Taken to museums and galleries as a child (%)	81	80	79	77	75	76
Total	322	198	526	500	500	1,000

Amongst the sample of visitors to Tate Liverpool, only 6% were visiting with children. On average, visitors spent one hour and ten minutes on site. 22% of visitors claimed to have made a donation, which was around £3.30 on average. The three most cited reasons for visiting Tate Liverpool were found to be the main collections (33%), the special paid temporary exhibitions (29%) and convenience/to kill time (11% each) (Table 10.3).

The vast majority of those in both surveys who had visited Tate Liverpool in the last 12 months had enjoyed their visit. 27% of visitors were familiar with the information on the work Tate Liverpool does inside the gallery. Only a minority of respondents in the general population survey said they were very or extremely familiar with the basic information presented on the work of Tate Liverpool inside the gallery. As might be expected, respondents in both surveys were more familiar with the work of Tate Liverpool inside the gallery than with its outreach work in the wider community (Table 10.2).

Using postcode data provide by respondents we were able to split the sample between those from the Northwest of England and those from the rest of the UK (excluding Northern Ireland). In the visitor's survey, 81% of Northwest visitors enjoyed their visit, compared to 75% of visitors from the rest of the UK. A significantly greater percentage of Northwest visitors (34%) were familiar with the information the work Tate Liverpool does inside the gallery, and with Tate Liverpool's wider community work outside the gallery (13%), compared to 15% and 4% respectively from visitors from the rest of the UK.

In the general population survey, significantly more respondents from the Northwest of England had visited TL at least once in their lifetime (36%), compared to the rest of the UK (14%). There was also significant and large difference in the percentage who had visited in the last 12 months, with 14% of those from the Northwest stating that they had visited Tate Liverpool in the last 12 months compared to only 6% from the rest of the UK. Tate Liverpool had 560,855 visitors in 2014. This equates to just under 1% of the UK population.<sup>14</sup> This suggests some element of overstatement in the general population online survey. This overstatement was especially high for respondents from the Northwest. This arguably may have been caused by the framing of the survey and the association between the cultural institution under study and factors like civic pride, personal interest/use, and tactical answers (if the respondent suspected that they would be filtered out of the survey for answering that they had not visited in the past 12 months).

<sup>14</sup> Population of UK excluding Northern Ireland is 62.3 million according to ONS mid-2013 estimates:

[www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk--england-and-wales--scotland-and-northern-ireland/2013/index.html](http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk--england-and-wales--scotland-and-northern-ireland/2013/index.html)

In the online survey, familiarity with Tate Liverpool was similar between the Northwest and rest of UK sample. Counterintuitively, familiarity with the work of Tate Liverpool in the wider community was significantly higher among the rest of the UK population than those from the Northwest. This again calls into question the accuracy of the responses to these particular questions.

**Table 10.3: Reasons for visiting the gallery (visitor survey)**

Reason	Northwest (%)	Rest of UK (%)	Whole survey (%)
Convenience of location, accessibility	8	16	11
To enjoy the main collections	30	38	33
To enjoy the gallery's special paid temporary exhibitions	35	21	29
To attend a talk/workshop	0.2	0	0.2
To entertain and/or educate children	0.6	0	0.6
To spend time with family or friends	10	7	9
To visit the café or the shop	4	3	4
To kill time/something to do	9	13	11
Other	3	3	3
Total	321	198	525

### 10.3. Attitudes

Visitors' general attitudes towards the arts and culture are depicted in Table 10.4. Some 17% of respondents in the general population survey and 27% in the visitors survey ranked heritage, arts, museums and culture amongst their three top priorities for public spending (from a list that included areas such as education, health, pensions, housing, environment, safety, transport, national defence and the economy). A majority of respondents agreed or strongly agreed that Tate Liverpool was a national treasure that should be protected for future generations, and had a value even for those people who do not visit. Only around a quarter of the general population surveyed agreed or strongly agreed that there were more important things to spend money on than protecting UK galleries such as Tate Liverpool.

Attitudes towards the arts and culture did not differ significantly between the Northwest and rest of the UK in either the visitor or general population surveys. Intuitively, in the general population survey, significantly more Northwest respondents agreed or strongly agreed that Tate Liverpool was a national treasure to be preserved for future generations (70%) compared with the rest of the UK (60%), and that Tate Liverpool has a value even for those people who do not visit (60%) compared with the rest of the UK (51%). Significantly more Northwest respondents disagreed or strongly disagreed that there were more important things to spend money on than protecting UK galleries (33%) compared with the rest of the UK (25%).

**Table 10.4: Attitudes towards culture and Tate Liverpool**

	Visitor survey			General population survey		
	Northwest	Rest of UK	Whole survey	Northwest	Rest of UK	Whole survey
Heritage, arts, museums and culture amongst the 3 top priorities for public spending (%)	26	28	27*	17	17	17*
Tate Liverpool is a national treasure to be preserved for future generations (% agree or strongly agree)	94	93	94*	70*	60*	65*
There are more important things to spend money on than protecting UK galleries such as Tate Liverpool (% disagree or strongly disagree)	52	46	21*	33*	25*	27*
Tate Liverpool has a value even for those people who do not visit (% agree or strongly agree)	73	70	73*	60*	51*	56*
Table	322	198	526	500	500	1000

Legend: \*  $p < 0.05$  significant difference between surveys (*t*-test)

## 10.4. Contingent valuation analysis

### 10.4.1. Summary willingness to pay measures

Table 10.5 shows those who are willing to pay something in principle to support the work of Tate Liverpool. In the visitor survey, 80% of respondents said they would in principle (or maybe would) be willing to make a one-off donation to support the work that Tate Liverpool does inside the gallery to improve visitor experience; 60% of respondents said they would (or maybe would) in principle be willing to pay a one off donation to support Tate Liverpool's community outreach work; while in the online general population survey, 47% of respondents said they would (or maybe would) be willing to make an annual donation, even if only a small amount, to help maintain and improve the gallery and to support the work that Tate Liverpool makes to the wider community (When asked the actual WTP question, and faced with a list of possible amounts, a total of 3% of visitors saying they were willing to make a donation to support the work of Tate Liverpool inside the gallery and 1% of visitors saying they were willing to make a donation to support the work of Tate Liverpool outside the gallery were found to have a zero willingness to pay, while a total of 6% of the general population sample expressed a zero WTP).

**Table 10.5. Willingness to pay**

	Work inside the gallery (Visitor survey)		Work outside the gallery (Visitor survey)		General population survey (inside & outside)	
	N.	%	N.	%	N.	%
Yes	272	51.71	138	29.61	105	10.5
Maybe	149	28.33	141	30.26	367	36.7
No	105	19.96	187	40.13	528	52.8
Total	526	100	466	100	1,000	100

Table 10.6 shows descriptive statistics for each WTP measure, including mean and median WTP amount. In the visitor survey, the mean willingness to pay to support the work that Tate Liverpool does inside the gallery was found to be £10.83. Median WTP was £5.50. These are **plausible donation values, in line with the prices currently charged for paid exhibitions in UK museums** (Note: for Tate Liverpool's Transmitting Andy Warhol exhibition, which was open from 7 November 2014 – 8 February 2015, overlapping with the fieldwork, adult entry was £8.80 (£8.00 without donation)).

**Table 10.6. Mean and median WTP**

Survey	WTP variable	N	Low 95%	Mean (£)	High 95%	Median (£)	Max (£)	Zeros (%)
Visitor	Donation: Work inside the gallery	517	9.15	10.83	12.51	5.50	250	22.4
Visitor	Donation: Work outside the gallery	462	6.58	8.00	9.42	2.50	100	41.1
Visitor	Visitor donation for work outside gallery, if agreed to pay donation for work inside	350	7.82	9.59*	11.37	5.50	100	30.0
Visitor	Visitor donation for work outside gallery, if refused to pay donation for work inside	111	1.37	3.05*	4.72	0	75	76.6
Visitor	Donation: Work inside the gallery (Northwest)	318	9.43	11.41	13.39	5.50	100	22.0
Visitor	Donation: Work inside the gallery (Rest of UK)	193	7.00	10.12	13.23	5.50	250	22.8
Visitor	Donation: Work outside the gallery (Northwest)	291	7.26	9.22*	11.18	5.50	100	35.7
Visitor	Donation: Work outside the gallery (Rest of UK)	165	4.14	6.10*	8.07	1.50	100	49.7
Online	Annual donation	1000	5.15	6.10	7.05	0	175	55.7
Online	Annual donation (Visitors)	101	7.35	12.34*	17.33	4.75	175	24.8
Online	Annual donation (Non-visitors)	888	4.51	5.41*	6.31	0	175	59.6
Online	Annual donation (Northwest)	500	5.12	6.37	7.62	0	112.5	52.8
Online	Annual donation (Rest of UK)	500	4.39	5.83	7.27	0	175	58.6

Legend: \*  $p < 0.05$  significant difference between survey subsamples (*t*-test)

Visitor mean willingness to pay a donation to support Tate Liverpool's community outreach work was found to be £8.00. As a proportion, this was around 75% of the price people were willing to pay to support the work that Tate Liverpool does inside the gallery. Median WTP was £2.50. As expected, mean willingness to pay to support Tate Liverpool's community outreach work was more than three times greater for those that had already agreed to pay a donation (£9.59) than for those who had refused (£3.05). This difference in WTP was significant at the 5% level. Similarly, 77% of those who had refused to pay a donation to support the work of Tate Liverpool inside the gallery provided zero value donation responses to support the work of TL outside of the gallery, compared with only 30% of those who had agreed to pay to support the work inside the gallery (Table 10.6).

The distributions of willingness to pay values for each payment mechanism are available in Annex G1.

We provide additional information for WTP broken down by key demographic groupings (Table 10.7). In the visitor survey, mean WTP for the work of TL inside the gallery was significantly higher for males (£13.20) and the employed (£12.11). Mean WTP for the work of TL in the wider community was significantly higher for males (£9.75) and those with higher levels of income (£9.91). This conforms to our expectation that those with higher spending capability should generally be more willing to make a higher donation. People living in the Northwest of England had significantly higher WTP for the work TL does in the wider community (£9.22). It is possible that high visitor WTP for the work of TL outside could be driven by the high number of local visitors. This would suggest that non-use values like local identity, pride, and community benefit are acting as drivers of visitor WTP.

In the general population survey, mean WTP was higher for those who claimed to have visited Tate Liverpool in the last 12 months (£12.34) than those who said they had not visited (£5.14). This difference was significant at the 5% level. Mean WTP was significantly higher for those aged under 50 (£7.59), males (£7.15), those with dependent children (£10.12), those with higher

educational attainment (£7.87), those with higher levels of income (£8.81), and the employed (£7.55). Mean WTP was slightly higher for those who lived in the Northwest (£6.37) than those who lived in the rest of the UK (excluding Northern Ireland) (£5.83), although this difference was not significant at the 5% level (Table 10.7).

**Table 10.7. Mean WTP by different socioeconomic groups (sample size in brackets)**

Socio-demographic characteristics	Visitor (use) WTP: Inside gallery	Visitor (use) WTP: Outside gallery	General population mean WTP
Age: Under 50	£9.76 (272)	£8.63 (244)	£7.59* (538)
Age: Above 50	£12.01 (245)	£7.30 (218)	£4.36* (462)
Gender: Female	£9.19* (306)	£6.82* (276)	£5.09* (508)
Gender: Male	£13.20* (211)	£9.75* (186)	£7.15* (492)
Parental status: Dependent children	£9.90 (71)	£8.46 (63)	£10.12* (300)
Parental status: No dependent children	£11.05 (441)	£7.98 (394)	£4.38* (700)
Education: Degree and above	£10.93 (349)	£8.52 (312)	£7.87* (346)
Education: Up to degree level	£10.67 (167)	£6.95 (149)	£5.16* (654)
Income: >£45,000 per annum (general pop median)	£12.34 (221)	£9.91* (193)	£8.81* (327)
Income: <£45,000 per annum	£10.12 (247)	£6.60* (224)	£4.78* (673)
Employed (including self-employed and part-time)	£12.11* (331)	£9.25* (292)	£7.55* (516)
Unemployed	£4.67 (6)	£3.83 (6)	£6.30 (68)
Student	£7.57 (66)	£6.26 (62)	£4.52 (42)
Retired	£9.40 (108)	£5.90 (97)	£3.64 (247)
Inactive/unpaid family worker	£5.00 (5)	£2.75 (4)	£5.43 (127)
Visited TL in the last 12 months	-	-	£12.34* (101)
Not visited TL in the last 12 months	-	-	£5.41* (888)
Northwest	11.41 (318)	£9.22* (291)	£6.37 (500)
Rest of UK (excluding Northern Ireland)	£10.12 (193)	£6.10 (165)	£5.83 (500)
London	£10.58 (20)	£13.24 (17)	-
East England	£5.17 (15)	£4.63 (15)	-
East Midlands	£8.25 (20)	£4.26 (17)	-
Northeast	£10.00 (5)	£4.13 (4)	-
Southeast	£4.30 (15)	£6.23 (15)	-
Southwest	£7.18 (14)	£2.75 (12)	-
West Midlands	£23.23 (22)	£7.64 (21)	-
Yorkshire & Humber	£8.70 (42)	£5.10 (31)	-
Scotland	£11.44 (18)	£8.43 (14)	-
Wales	£9.29 (19)	£2.91 (17)	-
Mean WTP	£10.83 (517)	£8.00 (462)	£6.10 (1000)

Legend: \*  $p < 0.05$  significant difference between survey subsamples (*t*-test)

## 10.4.2 Validity tests

In this section we assess the validity of the WTP findings and provide supporting evidence that the WTP results can be interpreted as reflecting the welfare changes associated with visiting Tate Liverpool and with the range of services it provides. There are a number of ways to test the validity of the results. Here, we discuss content validity, theoretical validity and the bias reduction mechanisms included in the survey design (specifically to tackle the issue of hypothetical bias).

### 10.4.2.1 Content validity

Content validity tests look at the adequacy, realism and neutrality of the survey instrument as well as at respondents' understanding, perception and reactions to the questionnaire. Additionally, the rate of protest provides valuable information on how respondents reacted to the scenarios.

We conducted stakeholder meetings at the project inception stage with Tate Liverpool's Director and senior management team to agree the valuation scenarios for the contingent valuations (for both the visitor and online surveys). We then undertook extensive testing of both draft survey instruments. We used a face-to-face pilot survey followed by in-depth cognitive debriefing about all parts of the questionnaire with 34 visitors at Tate Liverpool; and an online pilot survey with some cognitive follow-ups on key parts of the questionnaire with 70 panel respondents, in both cases mimicking the conditions in which the final surveys would be implemented.

On the visitor pilot survey, in terms of difficulty, 93% of respondents found the survey to be 'ok' or 'easy', with none describing it as 'very hard'. Roughly 75% didn't think there were any questions that were too sensitive. Finally, the cognitive debriefing showed that 52% of respondents found the survey length to be 'ok', while 48% found it 'a little long'. As a result the final questionnaire was substantively shortened and some of the longer questions dropped or edited down.

Pilot respondents were asked for their views on key and potentially complex parts of the questionnaire, such as the valuation and the wellbeing vignette questions. In terms of the valuation, some 92% of the sample found the scenario of government cuts to be 'realistic' or 'quite likely'. 79% said they understood the alternative valuation scenarios of donations to support the work Tate Liverpool does inside and outside the gallery fully. When asked about their certainty of paying the amount stated, 76% said they would definitely pay, and 24% said they would most likely pay.

As regards the online general population survey, the cognitive debriefing showed that 81% of respondents found the survey length to be 'ok'. Only 1.5% found it to be 'very long'. In terms of difficulty, 97.1% of respondents found the survey to be 'ok' or 'easy', with no respondents describing it as 'very hard'. Roughly 88% did not think that the survey used unclear or unusual language and 96% didn't think there were any questions that they felt uncomfortable answering. Moreover, 90% of the pilot sample agreed that the survey provided them with sufficient information.

In terms of the valuation section, 87% of respondents agreed that the survey provided enough information about the work of Tate Liverpool to answer the CV questions, and 80% found Tate Liverpool photos used in the survey instrument helpful for picturing its work and activities. Some 73% of the sample found the scenario of government cuts to be 'realistic' or 'quite likely'. Crucially, 86% of the sample did not find it hard to select a monetary value that they would be willing to pay as a donation to maintain Tate Liverpool and its work (and the remaining 14% found it only 'a little hard'). In addition, 76% stated that they fully understood the WTA scenario. In terms of the wellbeing vignette, 85% of the sample claimed to have understood it fully. Only 15% and 20% of respondents found the vignette question difficult or too long, respectively.

As with the NHM case study, these findings indicate that our pilot surveys instrument were generally well received by the public and were not on the whole difficult to understand. Only minor changes were deemed to be required, mostly aimed at shortening the visitor survey.

Finally, in terms of protest answers, follow-up questions to the WTP scenarios showed that only a very small proportion of respondents could be classified as protesters, i.e. stating a zero WTP when in fact they might value the change being proposed. In the visitor survey, only thirty respondents (less than 6% of the total) stated a zero WTP to the donation to support the work of TL inside the gallery because they either didn't agree with the donation system or they thought they paid enough for museums already through taxes. In the general population survey, only 86 people (less than 9% of the total) gave possible protest reasons for not offering a donation (e.g. that they paid enough already through other means such as taxes, and that Tate Liverpool should charge entrance fees rather than ask for donations).

### 10.4.2.2 Theoretical validity

As discussed in section 6.4.2.2, the credibility of estimated values from CV studies is commonly assessed by examining their theoretical validity – whether the relationship between WTP and other indicators are in accordance with expectations.

The set of variables included in the models matches that used in the NHM case study and is presented in Table 10.8.

**Table 10.8. Determinants of WTP**

Variable	Interpretation	Mean
Gender (male)	Dummy variable (1-male; 0-female)	0.46
Age (log)	Log of age (mid-point of intervals)	3.79
Education (university)	Dummy variable (1-university degree or higher; 0-otherwise)	0.46
Income (log)	Log of annual household income before tax (mid-point of intervals)	3.23
Children	Dummy variable (1-dependent children under 16; 0-otherwise)	0.26
Distance travelled/distance to TL (log)	Log of geographical distance from respondent postcode to TL (km)	32.43
Northwest resident (online survey only)	Dummy variable (1-Northwest resident; 0-otherwise)	0.54
Interaction: Northwest x Distance (online survey only)	Heterogeneous effect of distance and living in northwest (1-Northwest resident distance to TL; 0-otherwise)	6.52
Weekend (visitor survey only)	Dummy variable (1-visit during weekend; 0-otherwise)	0.16
Visits to TL in the last year	Continuous	0.94
Likely to visit TL in future	Likert scale 1-5 where 1 is 'not at all likely' and 5 is 'very likely'	2.90
Familiarity with TL	Likert scale 1-5 where 1 is 'not at all' familiar, and 5 is 'very familiar'	2.36
Familiarity with TL community work	Likert scale 1-5 where 1 is 'not at all' familiar, and 5 is 'very familiar'	1.91
Recently attended other galleries	Dummy variable (1-yes; 0-no)	0.65
Membership	Dummy variable (1-a member of a heritage, conservation, environmental or other organisation; 0-member of no organisation)	0.25
Closure of TL would reduce my life satisfaction	Dummy variable (1-closure of TL would reduce my life satisfaction; 0-closure of TL would have no effect on my satisfaction)	0.19
Agreement: 'Tate Liverpool is a national treasure to be preserved for future generations'	Likert scale 1-5 where 1 is 'strongly disagree' and 5 is 'strongly agree'	4.01
Agreement: 'Tate Liverpool has a value to non-visitors'	Likert scale 1-5 where 1 is 'strongly disagree' and 5 is 'strongly agree'	3.66

Further descriptions of all of the variables used in the TL case study can be found in Annex F.

#### (i) Visitor WTP to support work of Tate Liverpool inside the gallery (one-off donation)

Table 10.9 shows the results for Models 1, 2 and 3 for gallery visitors' willingness to pay a one off donation to support the work of TL inside the gallery, controlling for a range of factors.

Income is, as expected, a statistically significant and positive driver of higher willingness to pay across all three models (Bateman et al., 2002; Carson, 2012; Mourato et al., 2002).

Gender (male) is a significant positive driver of WTP in Models 1 and 3. Distance travelled is a significant and negative driver of willingness to pay in Model 1, suggesting perhaps that those who have outlaid the expense required to travel further to Tate

Liverpool have restrained budgets for donations, reducing their WTP. We note that this finding is the opposite of the distance effect found in the NHM case study. Since a larger proportion of visitors to TL originated in the Northwest region, it may be that we are capturing the true travel cost of visiting TL, and its effect as a budget constraint on WTP, as anticipated by the literature. In contrast, the NHM, as an international cultural institution based in London, may conceivably have a positive travel distance effect capturing the 'special' nature of a trip to London. The distance variable loses its statistical significance in Models 2 and 3, however, when experience and attitudinal variables are added. No behavioural variables are found to be significant drivers of WTP for the work of TL inside the gallery in Models 2 and 3.

Agreement with the statement that closure of Tate Liverpool for one year would reduce life satisfaction is a significant positive driver of WTP. This suggests that perceived personal effect of TL is an important factor in the value visitors place in the institution, which again conforms to prior expectations.

**Table 10.9. Drivers of visitor willingness to pay for work of TL inside the gallery (one-off donation)**

Variable	Model 1 Coefficient	Model 2 Coefficient	Model 3 Coefficient
Gender (male)	2.909*	2.819	3.091*
Age (log)	-0.678	-1.889	-2.938
Education (university)	0.839	0.467	0.095
Income (log)	2.546***	2.305**	2.822***
Children	-1.329	-1.3	-1.074
Distance travelled (log)	-0.011*	-0.009	-0.006
Weekend		1.938	2.221
Visits to TL in the last year		0.193	-0.036
Familiarity with TL		0.724	0.318
Recently attended other galleries		-2.154	-2.121
Membership		2.451	2.388
Closure of TL would reduce my life satisfaction			7.076***
Agreement: 'Tate Liverpool is a national treasure to be preserved for future generations'			1.21
Constant	3.626	6.942	3.627
Observations	502	498	496
r <sup>2</sup>	0.021	0.032	0.058

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at <10%. Reference group: (i) for gender ref = female; (ii) for education ref = all qualifications under Degree; (iii) for children ref = no children; (iv) for weekend ref = weekday; (v) for membership ref = membership of no organisation. Sample is restricted to England, Wales and Scotland and those aged 16 and over. Heteroskedasticity-robust standard errors.

### **(ii) Visitor willingness to donate towards TL's community work (additional one-off donation over and above donation to support work of TL inside the gallery)**

Table 10.10 shows the results for Models 1, 2 and 3 for museum visitors' willingness to pay a donation (over and above donation to support work of TL inside the gallery) towards community and outreach work, controlling for a range of factors.

Income is once again a significant and positive driver of willingness to make a donation in Models 2 and 3. This suggests that income is a driver of willingness to donate towards the work of Tate Liverpool outside the gallery. Age is a significant negative driver of WTP in Model 3. Distance travelled is again a negative determinant of WTP, albeit significantly only in Model 1.

In Models 2 and 3, having recently attended other galleries turns out to be a significant positive determinant of WTP. Although this sits oddly with previous findings suggesting that knowledge and experience of alternative sources of cultural value reduces WTP for the institution under study, it is consistent with recent attendance at other galleries being a general indicator of preferences for cultural engagement (and especially given that here we are valuing TL's external outreach work).

Visiting Tate Liverpool at the weekend is also found to be a statistically significant positive driver of WTP, at least in Model 3. This may possibly be capturing the 'special' effect of visiting TL as part of a weekend day out in Liverpool. Conceivably, it may be that those who attend the gallery during non-work hours derive more benefit than those who visit during the week, say because the experience is heightened in contrast to activities performed during the working week, or because the visit is part of a wider leisure experience in the city. However, in those cases we might have expected a statistically significant result in WTP for work inside the gallery too, which is not what we see in Table 10.9.

Agreement with the statement that closure of Tate Liverpool for one year would reduce life satisfaction is again a significant positive driver of WTP a donation for the work of TL.

**Table 10.10. Drivers of visitor willingness to make a one-off donation for TL's work outside the gallery (on top of donation for work inside gallery)**

Variable	Model 1 Coefficient	Model 2 Coefficient	Model 3 Coefficient
Gender (male)	2.698	2.274	2.408
Age (log)	-2.516	-2.445	-3.313*
Education (university)	0.809	0.243	-0.135
Income (log)	2.372	2.537*	2.982**
Children	-0.19	-0.186	-0.175
Distance travelled (log)	-0.014*	-0.011	-0.008
Weekend		2.361	2.506*
Visits to TL in the last year		0.235	0.054
Familiarity with TL community work		0.872	0.573
Recently attended other galleries		3.119**	2.901**
Membership		-0.97	-0.958
Closure of TL would reduce my life satisfaction			5.981***
Agreement: 'Tate Liverpool has a value to non-visitors'			0.668
Constant	8.657	2.572	1.622
Observations	449	445	443
r <sup>2</sup>	0.027	0.044	0.069

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at <10%. Reference group: (i) for gender ref = female; (ii) for education ref = all qualifications under Degree; (iii) for children ref = no children; (iv) for weekend ref = weekday; (v) for membership ref = membership of no organisation. Sample is restricted to England, Wales and Scotland and those aged 16 and over. Heteroskedasticity-robust standard errors.

### (iii) General population willingness to donate towards the work of TL inside and outside the gallery (annual donation)

Table 10.11 shows the results for Models 1, 2 and 3 for the general population survey willingness to pay an annual donation towards the work of Tate Liverpool inside and outside the gallery, controlling for a range of factors.

Income is a significant positive driver of willingness to pay a donation in all models, consistent with theoretical expectations as, it turns out, is being male. Older age is a negative driver of WTP in Model 1 only. Having dependent children is a significant positive determinant of WTP in Models 1 and 2, suggesting that we may be (weakly) capturing an element of non-use bequest value, whereby individuals support the work of cultural institutions like Tate Liverpool because they value their continued existence for the benefit of their children.

Owing to co-linearity between living in the Northwest of England and distance to Tate Liverpool in the general population survey, we created a variable interacting the two. Living in the Northwest turns out to be a strong positive driver of WTP in

Models 1 and 2, perhaps reflecting considerations like local pride and public interest in the wider cultural status of the region. For the rest of the country, distance to TL is a significant positive driver of WTP in Models 1 and 2, suggesting, counterintuitively, that the *further away* a respondent is from TL, the *more* they value the institution. For the interaction variable of distance to TL within the Northwest, we find that distance is a significant negative driver of WTP in Model 1, suggesting that there is a weak effect whereby the closer a respondent lives to TL (i.e. in the Liverpool and Merseyside area), the more they value Tate Liverpool. This conforms with our expectation that people living within the local Merseyside region value Tate Liverpool more than those in other parts of the Northwest and, other things equal, those in the rest of the country too.

Familiarity with TL's community work is a significant positive driver of WTP. This conforms to our expectations that individuals who are more aware of the work of a cultural institution will value it more highly. Having recently attended alternative galleries is again a significant positive determinant of WTP.

As in the case of the NHM case study, agreement with the statement that closure of Tate Liverpool for one year would reduce life satisfaction is a significant positive driver of willingness to make a donation for the work of TL in the general population survey. Agreement with the statement that Tate Liverpool has a value to non-users is also a significant and positive driver of WTP, direct confirmation that we are capturing an element of non-use values in our general population WTP.

**Table 10.11. Drivers of general population willingness to donate towards the work of TL inside and outside the gallery (annual donation)**

Variable	Model 1 Coefficient	Model 2 Coefficient	Model 3 Coefficient
Gender (male)	2.526**	1.911*	2.211**
Age (log)	-3.114**	-1.252	-1.39
Education (university)	1.511	0.665	1.157
Income (log)	2.712***	1.832**	1.776**
Children	4.088***	2.217*	2.045
Distance to TL (log)	2.703**	2.579**	1.605
Northwest	16.202**	11.487*	6.463
Interaction: Northwest x Distance	-2.933**	-1.964	-0.921
Visits to TL in the last year		0.215	-0.079
Likely to visit TL in future		0.351	-0.231
Familiarity with TL community work		2.891***	2.101***
Recently attended other galleries		2.230**	2.145**
Membership		1.682	1.757
Closure of TL would reduce my life satisfaction			5.640**
Agreement: 'Tate Liverpool has a value to non-visitors'			1.184**
Constant	-7.037	-17.872**	-14.403
Observations	902	860	847
r <sup>2</sup>	0.075	0.14	0.153

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at <10%. Reference group: (i) for gender ref = female; (ii) for education ref = all qualifications under Degree; (iii) for children ref = no children; (iv) for weekend ref = weekday; (v) for membership ref = membership of no organisation. Controls are included for experimental branches of cheap talk/oath script in online survey. Sample is restricted to England, Wales and Scotland and those aged 16 and over. Heteroskedasticity-robust standard errors.

Taken as a whole, these findings suggest that the WTP estimates in this case study have a high degree of theoretical validity.

#### 10.4.2.3 Hypothetical bias corrections

We included a number of mechanisms in the survey design to tackle the possibility of hypothetical bias: in particular, the inclusion of certainty measures following the donation questions, and the use of cheap talk and oath entreaties.

**(i) Certainty**

As noted earlier, in some published studies, the effect of certainty on WTP has been found to be negative, suggesting that respondents want to pay smaller amounts when they are more certain about their valuation (Bedate et al., 2009).

**Table 10.12. Impact of respondent certainty that they would pay the stated amount on willingness to pay to support the work of Tate Liverpool inside and outside the gallery**

Variable	Visitor survey: Use value (work inside gallery) Coefficient	General population survey Coefficient
Certainty (very certain)	-2.847	6.312**
Gender (male)	3.41	3.831**
Age (log)	-2.225	1.198
Education (university)	0.562	2.885
Income (log)	2.790**	3.494**
Children	-1.781	2.613
Distance travelled/distance to TL (log)	-0.002	3.359
Northwest resident	-	14.015
Interaction: Northwest x distance	-	-1.848
Weekend	2.237	-
Visits to TL in the last year	-0.009	0.226
Likely to visit TL in future	-	-1.402
Familiarity with TL	1.252	-
Familiarity with TL community work	-	1.486
Recently attended other galleries	-1.867	4.063*
Membership	4.211*	1.228
Closure of TL would reduce my life satisfaction	6.819**	4.363
Agreement: 'Tate Liverpool is a national treasure to be preserved for future generations'	0.941	-
Agreement: 'Tate Liverpool has a value to non-visitors'	-	-0.333
Public spending on heritage, arts, museums and culture as priority	0.701	-3.246
Constant	1.696	-27.631
Observations	394	400
r <sup>2</sup>	0.072	0.12

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at <10%. Reference group: (i) for gender ref = female; (ii) for education ref = all qualifications under Degree; (iii) for children ref = no children; (iv) for weekend ref = weekday; (v) for membership ref = membership of no organisation. Controls are included for experimental branches of cheap talk/oath script in online survey. Sample is restricted to England, Wales and Scotland and those aged 16 and over. Heteroskedasticity-robust standard errors.

Table 10.12 shows certainty as a driver of willingness to pay only for those that answered that they would (or maybe would) be willing to pay a donation to support the work of TL. In the online survey, we find that the level of certainty turns out to be a significantly *positive* determinant of the stated donation. This might be explained if those responding with a large amount have thought more deeply about the true value of the change being proposed, and that their higher value responses are therefore more considered (more certain) than the lower value responses of others. That said, no significant association in either

direction is found between certainty and WTP in the visitor survey, perhaps because the proportion of those who claimed to be somewhat or very certain was even higher than in the general population survey, meaning that there is very little variation in the sample (Table 10.13).

**Table 10.13. Level of respondent certainty of their actual willingness to pay to support the work of Tate Liverpool inside and outside the gallery**

Level of certainty	Visitor survey (%)	General population survey (%)	Whole sample (%)
Not certain	5.54	14.83	10.48
Somewhat certain	54.46	63.98	59.53
Very certain	40.00	21.19	29.99
Observations	415	472	887

Note: Respondents who answered No to the dichotomous WTP question were not asked the certainty of their response.

### (ii) Cheap talk and oath scripts entreaties

Table 10.14 shows that the oath script and cheap talk script when used in combination had a significant negative effect on stated willingness to pay in the general population. That is, the presence of a promise to answer the questions as truthfully as possible along with a cheap talk script, by reminding respondents not to provide incredible estimates, may have succeeded in containing biases. On their own, we find that neither the oath script nor the cheap talk had a significant effect on stated WTP, in contrast to the NHM case study where the oath script appeared to have some impact. Clearly, since the real value of donations is not observable, we cannot prove that the value obtained with the combined entreaties is any closer to the true WTP than the value without entreaties, but the results are suggestive that that is the case.

**Table 10.14. Entreaties experiments**

Variable	Coefficient	Variable	Coefficient
Oaths script and cheap talk	-2.399**	Cheap talk only	-0.432
Oath script only	0.048	Constant	-14.407
Observations	847	r <sup>2</sup>	0.153

Notes: \*\*\* significance at <1%; \*\* significance at <5%; \* significance at <10%. Reference = no cheap talk or oath script. Sample is restricted to England, Wales and Scotland and those aged 16 and over. OLS regressions, controlled for a range of socio-demographic and experience factors (not shown). Heteroskedasticity-robust standard errors.

## 10.5. Wellbeing analysis and valuation

### 10.5.1. Wellbeing summary statistics

**Table 10.15. Mean wellbeing scores across the visitor and online surveys**

	Mean life satisfaction	Mean happiness (last hour)	Mean purpose (last hour)
<b>Overall</b>	6.946	7.102	6.928
<b>Visitors</b>	7.830	8.132	8.095
<b>General population</b>	6.486	6.563	6.316

As in the NHM case study, all reported measures of wellbeing were on average higher for visitors than respondents to the general population survey (Table 10.15).

**Table 10.16. Mean scores for happiness and purpose in the last hour by different activities at Tate Liverpool**

Activity	Mean happiness (last hour)	Mean purpose (last hour)
Regular exhibition/gallery	7.952	7.849
Paid exhibition/gallery	8.465	8.527
Lectures, talks, or workshops	8.4	9.4
Tate Liverpool cafe	7.722	7.944
Relaxing in non-exhibition area	9	8.133
Other	7.947	8.316

Relaxing in a non-exhibition area of the gallery turned out to be associated with the highest levels of happiness, followed by going to a paid exhibition. For purpose, going to lectures, talks or workshops was the highest ranking activity followed by going to a paid exhibition. Relaxing in the gallery was associated with high levels of happiness, but was not as purposeful, whilst going to lectures, talks or workshops was very purposeful and average for happiness.

### 10.5.2. Momentary wellbeing analysis

Tables 10.17 and 10.18 show the results for estimating equation (8) for happiness and purpose in the last hour using the data collected in the TL visitor survey. Note that the number of observations used in the model is lower than the overall number of survey responses due to non-response in some of the variables (this caveat applies to all of the models estimated).

After controlling for a range of socio-demographic factors as well as current levels of life satisfaction, activities at Tate Liverpool rank highly in terms of associations with happiness. Not many of the general activities are found to be statistically significant, but after 'playing sports', 'relaxing in a non-exhibition area of the gallery' and 'going to a paid exhibition in Tate Liverpool' rank as the second and third highest activities for happiness.

**Table 10.17 Association between activities and happiness in the last hour**

Tate Liverpool	Coefficient	Tate Liverpool	Coefficient
Regular exhibition/gallery	0.456**	Paid exhibition/gallery	0.712***
Lectures, talks, or workshops	-0.169	Tate Liverpool café	0.223
Relaxing in non-exhibition area	1.089***	Other	0.723
Other activities	Coefficient	Other activities	Coefficient
Internet	-0.242	Care adults	0.02
Care children	0.02	Computer games	-0.085
Cultural activities	0.893	Eating	-0.259
Hobbies	0.201	Housework	-0.062
Reading/TV	0.163	Sick in bed	-2.168**
Sleeping	0.148	Sports	1.273***
Talking	0.032	Travel	0.194
Cooking	0.255	Music	0.681**
Something else	-0.158	Observations	1463
r <sup>2</sup>	0.624		

Legend: \*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ . Activity refers to the main activity in the last hour. The following control variables included in the model but not shown here: gender, health, income, marital status, education, parental status, employment status, social networks, life satisfaction. Reference case for activities is working/studying. Heteroskedasticity-robust standard errors.

After controlling for a range of socio-demographic factors, as well as current levels of life satisfaction, activities at Tate Liverpool also rank highly in terms of associations with purpose (Table 10.18). All activities at Tate Liverpool, except for relaxing in a non-exhibition area of the gallery, are positively associated with purpose. In the following order – 'other activities at Tate Liverpool'; 'going to lectures, talks and workshops'; going to paid exhibitions'; and 'spending time in the café' – are the highest ranking activities for purpose. The highest-ranking non-TL activity is 'listening to music'.

**Table 10.18. Association between activities and purpose in the last hour**

Tate Liverpool	Coefficient	Tate Liverpool	Coefficient
Regular exhibition/gallery	0.662***	Paid exhibition/gallery	1.041***
Lectures, talks, or workshops	1.234***	Tate Liverpool café	0.957**
Relaxing in non-exhibition area	0.216	Other	1.533***
Other activities	Coefficient	Other activities	Coefficient
Internet	-0.626***	Care adults	0.63
Care children	0.533	Computer games	-0.362
Cultural activities	0.779	Eating	-0.555*
Hobbies	0.742*	Housework	-0.101
Reading/TV	-0.468*	Sick in bed	-1.752
Sleeping	-0.056	Sports	0.844
Talking	0.104	Travel	0.254
Cooking	0.256	Music	0.823**
Something else	-0.271	Observations	1463
r <sup>2</sup>	0.498		

Legend: \*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$  Activity refers to the main activity in the last hour. The following control variables included in the model but not shown here: gender, health, income, marital status, education, parental status, employment status, social networks, life satisfaction. Reference case for activities is working/studying. Heteroskedasticity-robust standard errors.

The results echo the NHM case study in providing further evidence that cultural activities are strongly associated with higher levels of wellbeing in the moment: people are happier and have a higher sense of purpose when they engage in cultural activities.

### 10.5.3. Wellbeing valuation analysis

#### (i) Standard wellbeing valuation approach

Table 10.19 shows the results from estimating equation (9). This is the association between engagement with Tate Liverpool and life satisfaction after controlling for the main determinants of wellbeing. We look at the effect of being a visitor at Tate Liverpool over the past 12 months – i.e., 'Visited Tate' is a binary variable that equals 1 if the individual visited Tate Liverpool at least once over the past 12 months.

Visiting Tate Liverpool has a positive but statistically insignificant association with life satisfaction after controlling for a wide range of socio-demographic variables. But being a member of Tate Liverpool has a large significant positive association with life satisfaction.

**Table 10.19. Association between visiting Tate Liverpool (any number of times) and life satisfaction**

Museums variables	Coefficient	Museums variables	Coefficient
Visited TL	0.11	Member TL	0.386*
Museums/galleries growing up	0.279	Training/classes arts	-0.167
Live near museum/gallery	0.054	Friends/family go museums/galleries	-0.377*
Control variables	Coefficient	Control variables	Coefficient
Online survey	-1.009***	Male	0.071
Age (20-29)	0.153	Age (30-39)	0.425
Age (40-49)	0.28	Age (50-59)	0.522*
Age (60-69)	0.688**	Age (over 70)	1.048***
Married	0.572***	Civil partner	0.295
Separated	-0.812*	Divorced	0.164
Widowed	0.103	GCSE	0.46
A level	0.221	Professional qualification	0.384
Degree	0.227	Higher degree	0.073
Self-employed	0.920***	Full-time	1.180***
Part-time	1.124***	Student	1.525***
Home care	0.915**	Retiree	1.562***
Health excellent	2.279***	Health very good	1.998***
Health good	1.457***	Health fair	0.904***
Income (£15-19k)	0.345*	Income (£20-29k)	0.302*
Income (£30-39k)	0.347**	Income (£40-49k)	0.421**
Income (£50-59k)	0.418**	Income (£60-79k)	0.461*
Income (£80-99k)	0.806***	Income (£100-149k)	-0.023
Income (£150k+)	-0.164	Children	-0.012
Socialises most days	0.205*	Socialises weekly	0.133
Constant	3.495***	Observations	1425
r <sup>2</sup>	0.304		

Notes: Model Legend: \*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ . Reference case is people who did not visit Tate Liverpool at all in the past 12 months. Heteroskedasticity-robust standard errors.

In Table 10.20 we use a continuous visit variable. In the simplest unconditional regression model, whereby life satisfaction is regressed on number of Tate visits (column 1 in Table 10.20), we find that visits to Tate Liverpool has a significant positive association with life satisfaction. However, once the main control variables have been included in the model, visits to Tate Liverpool become insignificant (column 2 in Table 10.20). We tested this relationship under different functional forms for the visit variable, using models with natural logs; squared terms and separate visit frequency categories (results not presented here). Visit frequency remains insignificant in all cases.

**Table 10.20. Association between Tate Liverpool visit frequency in the past year and life satisfaction**

Museums variables	Model 1 Coefficient	Model 2 Coefficient	Museums variables	Model 1 Coefficient	Model 2 Coefficient
TL visit frequency	0.162***	0.004	Member TL		0.394*
Museums/galleries growing up		0.282	Training/classes arts		-0.167
Live near museum/gallery		0.054	Friends/family go museums/ galleries		-0.379*
Control variables	Model 1 Coefficient	Model 2 Coefficient	Control variables	Model 1 Coefficient	Model 2 Coefficient
Online		-1.095***	Male		0.071
Age (20-29)		0.151	Age (30-39)		0.421
Age (40-49)		0.279	Age (50-59)		0.519*
Age (60-69)		0.685**	Age (over 70)		1.039***
Married		0.574***	Civil partner		0.297
Separated		-0.807*	Divorced		0.162
Widowed		0.101	GCSE		0.456
A level		0.217	Professional qualification		0.381
Degree		0.226	Higher degree		0.069
Self-employed		0.924***	Full-time		1.192***
Part-time		1.128***	Student		1.535***
Home care		0.917**	Retiree		1.569***
Health excellent		2.285***	Health very good		2.002***
Health good		1.459***	Health fair		0.904***
Income (£15-19k)		0.346*	Income (£20-29k)		0.300*
Income (£30-39k)		0.345**	Income (£40-49k)		0.419**
Income (£50-59k)		0.418**	Income (£60-79k)		0.460*
Income (£80-99k)		0.800***	Income (£100-149k)		-0.032
Income (£150k+)		-0.158	Children		-0.004
Socialises most days		0.210*	Socialises weekly		0.135
Constant	6.800***	3.583***	Observations	1514	1425
r <sup>2</sup>	0.024	0.304			

Legend: \*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ . Reference case is people who did not visit Tate Liverpool at all in the past 12 months. Heteroskedasticity-robust standard errors.

As with the NHM study, estimating the wellbeing effect of the frequency of visits to Tate Liverpool proves problematic. This time, we do not find the positive association between Tate Liverpool visits and life satisfaction to be significant. This finding, however, could conceivably reflect a number of factors other than the absence of a genuine effect, including measurement error (e.g., people misreporting the number of times they have actually visited TL), and the possibility that the sample is insufficiently large to allow a significant impact to be detected.

We now turn to the analysis of the hybrid contingent-wellbeing valuation results and vignette analysis to assess whether they provide additional methods for valuing the work of cultural institutions.

### (ii) Hybrid contingent-wellbeing valuation

Table 10.21 shows responses to the hybrid contingent-wellbeing valuation scenario in the TL case. In the visitor survey, 21% of respondents said that closure of Tate Liverpool would reduce their life satisfaction, compared with 17% in the general population survey. 79% of respondents to the visitor survey said that closure of Tate Liverpool would have no effect,<sup>15</sup> slightly lower than in the general population (81%). Less than 1% of respondents to the visitor survey said that the closure would actually *increase* their life satisfaction, compared with around 2% of the general population survey. We treated these responses as protest votes in our data and removed them from the sample.

**Table 10.21 Willingness to Accept closure of Tate Liverpool**

	Visitor survey		General population		Total	
	N	%	N	%	N	%
The closure would have no effect on my life satisfaction	415	78.9	811	81.1	1,226	80.34
The closure would reduce my life satisfaction	109	20.72	171	17.1	280	18.35
The closure would increase my life satisfaction	2	0.38	18	1.8	20	1.31
Total	526	100	1,000	100	1,526	100

Table 10.22 shows the average WTA for closure of Tate Liverpool for one year to keep life satisfaction constant. This question was asked only of people that said their life satisfaction would decrease due to the closure of Tate Liverpool.<sup>16</sup>

For the group of people that say their life satisfaction would be reduced due to the closure of Tate Liverpool the average level of WTA for closure of Tate Liverpool for one year turned out to be £31.83. However, there were also those people that said that the closure would have no effect on their life satisfaction. For consistency with the analysis in the NHM case study we assume that WTA closure of Tate Liverpool for one year was a nominal £1 for this group (which equates to about the bottom third of WTA values for the group that said their life satisfaction would decrease).

**For the full sample, the average level of WTA for closure of Tate Liverpool for one year works out as £6.55 (Figure 10.22).**<sup>17</sup> If we assume that the respondents would continue to visit Tate Liverpool at the same rate as they reported for the previous year then we can calculate a **value per Tate Liverpool visit, of £7.13 per visit** (based on an average annual visit frequency of 0.919 among respondents who answered the WTA question). This value can be compared to WTP for entry fees, but since this question was not asked in our contingent valuation survey there is no direct value that we can compare this WTA figure with. The WTA per visit value is, however, lower than the visitor donation for work inside the gallery (£10.83), and is not dissimilar to the per-visit value derived in the NHM study (£6.89). We tentatively conclude, therefore, that the use of the explicit reference to keeping life satisfaction constant seems again to have reduced the well-documented hypothetical biases in WTA figures.

<sup>15</sup> Note: in the TL survey, the wording of option one was changed from "The closure would have very little effect on my life satisfaction" as in the NHM survey to "The closure would have no effect on my life satisfaction" to reduce the potential for confusion among respondents.

<sup>16</sup> We note that, as expected, very few people (around 1% of the sample) said that their life satisfaction would increase due to the closure. In theory these people would have a (positive) WTP for the closure. However, visit frequency among this group was significantly higher than the general survey population and hence their responses to this question seem to be counterintuitive. Therefore, since this group of people make up a very small part of the sample we exclude them from the analysis.

<sup>17</sup> Note that using a conservative estimate where those who indicate that the closure would have no effect on their satisfaction are recorded as £0, we find a mean WTA of £5.65 for the full sample.

The approach allows us to look at the life satisfaction impact of additional visits to Tate Liverpool, which is not possible with the standard WV method. In the WTA scenario, people are explicitly made to think about the use value of their visits, as they are told that they are not able to visit the gallery at all for one year.

Whereas in the standard WV analysis, visit frequency has no effect on life satisfaction, in the WTA scenario people, it turns out, place a reasonable value on a Tate Liverpool visit. As the task is to compensate people for a loss in life satisfaction due to the gallery’s closure this implies that they actually predict a small impact on life satisfaction from Tate Liverpool visits. Echoing the NHM case study, this finding raises further questions as to whether the earlier lack of estimated impact of TL visit frequency on life satisfaction genuinely reflects the absence of a relationship.

As in the NHM study, we reverse engineer from the WTA value a predicted life satisfaction impact. That is, we replicate a coefficient for Tate Liverpool visits based on the WTA values, based on equation (10’):

$$\text{Value of Tate Liverpool engagement} = \beta_2 / \beta_1 \quad (10')$$

where we know the left hand side of the equation for a single visit (£7.13) and the impact of income ( $\beta_1$ ) from the income model in Annex C ( $\beta_1 = 1.1$ ). The value for  $\beta_2$  which solves this equation is 0.0005.<sup>18</sup> In other words, the WTA value together with the estimate for  $\beta_1$  from the income model suggests that people predict that not being able to visit Tate Liverpool once during the year would reduce their life satisfaction by 0.0005 points.

**Table 10.22. WTA values based on life satisfaction**

Mean WTA	Average visits	WTA (per visit)	Estimated LS coefficient (annual)	Estimated LS coefficient (visit)
£6.55	0.919	£7.13	0.0004	0.0005

**(iii) Vignette-based wellbeing valuation**

The vignette study can, in theory, also be used to derive estimates of the impact of TL visits on life satisfaction, and the value of these visits. However, we find that the life satisfaction scores for the vignettes involving higher TL visit frequency are no higher than the life satisfaction scores for real life, even though the vignettes envisaged, on average, 7.8 more visits per year.

This result is in marked contrast with what we find in the NHM study. It therefore puts into question the validity of using vignettes to assess changes in subjective wellbeing from engagement with a cultural institution. In particular, the vignette results appear to be inconsistent with: (i) our momentary wellbeing impacts, which show that a large number of activities at Tate Liverpool are in fact associated with higher levels of wellbeing, (ii) the life satisfaction analysis, where, although we did not get statistically significant results, the association between Tate visits and life satisfaction is positive, and (iii) the hybrid contingent-wellbeing valuation results, where we are able to discern a small positive impact of TL visits on predicted life satisfaction.

## 11. Case study two: Discussion and interpretation of the results

Table 11.1 attempts to bring to together all the key findings from the Tate Liverpool study. Care must be taken when using and interpreting the range of values estimated. We have estimated three core types of WTP with different meanings depending on the context they refer to.

Visitor use value provides an estimate of the direct use value to TL visitors, elicited through a visitor donation.

Visitor non-use value in this case captures the work of TL in the wider community and was elicited as a separate donation.

The online survey captured non-use and option values for Tate Liverpool as a whole, including its work inside and outside the gallery, amongst the general UK population (excluding Northern Ireland) as an annual donation. We note that in the general population survey the work of TL outside the gallery cannot be disentangled from its work inside the gallery. It is not, for example, possible to equate non-use value amongst visitors to non-use value in the general population.

The values elicited in Table 11.1 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

<sup>18</sup> Note that we use the non-linear version of equation (10’) from Annex C (i.e., equation (C3)) with a figure for sample median income of £25,000 pa.

or non-visitors, and among visitors whether we are interested in direct use or non-use values, as well as the influence that the payment mechanism (one-off or annual donation) has on the size of the value elicited.

**Table 11.1. Tate Liverpool valuation results**

Survey	Use/non-use value	Valuation variable	Mean	Median
Visitor	Use	Visitor donation for work inside the gallery (WTP)	£10.83	£5.50
Visitor	Non-use (visitors)	Visitor donation for work outside the gallery (WTP)	£8.00	£2.50
Online	Non-use and option (general population)	Annual donation for work inside and outside the gallery (WTP)	£6.10	£0
Visitor and online		Compensation for closure (per visit) (WTA) (hybrid contingent-wellbeing valuation)	£7.13	N/A
Visitor and online		Vignette wellbeing value (per visit)	N/A	N/A

*Note: values calculated using mid-point of intervals from the payment cards. Standard WV values are excluded due to the discussed problems concerning priming effects in the survey.*

A broad conclusion is that our study has produced **visitor WTP donation values to support the work of Tate Liverpool inside the gallery which are of a plausible magnitude, and in line with prices currently charged for paid exhibitions in UK museums.** Conforming to expectations, **visitor WTP amounts to support the work of Tate Liverpool inside the gallery are affected by income, and the perceived personal effect of TL closure.**

**Visitors' average levels of stated donations to support the work of Tate Liverpool outside of the gallery (over and above any previous donation) are also of a plausible magnitude, and driven by income, whether respondents had recently attended other galleries, and the perceived personal impacts of TL closure.**

**Mean and median annual donations by the general population to support the work of TL, inside and outside the gallery, are also of a credible order of magnitude.** The results accord again strongly with theoretically relevant factors. Income is again a positive determinant of WTP. **Familiarity with the work Tate Liverpool does outside the gallery in the community** is also a strong predictor. We also find evidence that **attitudes towards the non-use value of cultural institutions and attendance at other galleries determines willingness to pay.**

We note that visitor donations are considerably higher than stated donations elicited from the general population – average hypothetical visitor donations for both work inside and outside the gallery combined amount to roughly three times the equivalent for the general population. This difference is perhaps not so surprising: we would expect visitors to value art and culture more highly than the rest of the public.

We note that the mean WTP values we report are influenced by the presence of outliers at and above £100 (Annex G1). These have been retained in the analysis because the values are not implausible for support of a cultural institution, and analysis of socio-demographic background factors shows, intuitively, that they correspond to higher income groups. That said, **the median visitor WTP values, of £5.50 in donations to support the work of Tate Liverpool inside the gallery and £2.50 to support their work outside the gallery, may provide a better indication of what the 'typical' person is prepared to pay.** For the general population, the presence of a high percentage of zero responses drives down the median amount to £0.

Overall, we detected **few protest values** across the two surveys. There was some suggestion of hypothetical bias in the general population survey that used an annual donation payment mechanism, as, when combined, **the oath and cheap talk scripts** – which aim to reduce these types of biases – had a significant negative effect on WTP. Each of the oath and cheap talk scripts separately did not significantly affect WTP, however, though this might plausibly reflect the relatively small size of the samples. Moreover, all of the survey scripts maintained a basic budget reminder, which would already have a constraining effect on WTP. Under these circumstances, it is perhaps not surprising therefore that the additional impact of the scripts on their own was small.

These results suggest that **cheap talk combined with oath scripts** is another promising way to address hypothetical bias for valuations in the cultural sector.

As in the NHM case study, in order to help address problems in valuation surveys that arise from informational constraints, we used a host of **best-practice methods** including:

- Use of photos and detailed descriptions of the good/service plus payment vehicle.
- Checking familiarity with the various pieces of information provided during test surveys.
- A wide payment range in the payment ladder (informed by the results of our survey testing) to reduce anchoring and range effects.

But, again as is the case in the NHM study, it is likely that the stated **WTP donations, particularly in the visitor survey, are likely to have been primed by the current donation suggestions made by the gallery**, which we can do little about, at least within the confines of the present CV. Tate Liverpool has three donation boxes in the gallery, including one at the exit where interviews were conducted, with a suggested donation of £4. Survey responses may also have been primed by the current entrance fees for paid exhibitions (which were around £8 in Tate Liverpool during the period when the fieldwork occurred). And in this sense, arguably, the plausibility of the responses to the WTP questions is not so surprising, as they are designed to mimic what would happen if we had a real market for gallery support via donations. We cannot therefore rule out that the contingent valuation might have delivered less plausible valuation estimates in cultural settings where there was no relevant market price to act as ‘anchor’ (as would be the case, for example, when an institution houses unique and iconic collections).

**To sum up, the case study of Tate Liverpool follows best-practice in CV methodology and finds that the method is a viable approach for measuring use and to some extent non-use (and option) value. The method produces realistic values that vary in ways predicted by theory and consistent with previous findings, across a range of scenarios, payment vehicles and population groups.**

Our analysis of the association between activities at Tate Liverpool and happiness and purpose at the momentary level shows that **experiences within the gallery and partaking in cultural activities and events at Tate Liverpool are associated with significantly higher levels of wellbeing than other activities**. These results are in line with previous findings using ESM data that show that cultural activities rank very highly in terms of impacts on SWB. This provides some **rationale for why people are willing to pay to visit Tate Liverpool and we would expect there to be an association between TL engagement and SWB over the longer term**.

The analysis of the life satisfaction data, however, does not find the positive association between Tate Liverpool visits over the past 12 months and life satisfaction to be statistically significant. As with the NHM study, detecting the effect of visits to Tate Liverpool on non-momentary wellbeing proves problematic. This suggests that **although previous research has shown WV to work well with large national datasets, implementing it in the case of individual cultural institutions may be less appropriate**.

**We also find that the vignettes approach does not provide plausible results in the case of Tate Liverpool (unlike in the NHM study). The lack of consistent findings warrants further investigation. Although tested across only two case studies, it raises doubts about the validity of currently using vignettes to assess the impact of changes in engagement with a single cultural institution on wellbeing.**

What we call the **hybrid contingent-wellbeing valuation approach, i.e. the WTA scenario based on life satisfaction, for a scenario involving a hypothetical closure of TL for one year, again provides plausible values per visit**. This hybrid method offers an approach to valuing cultural services on a per visit basis without the need to set out a WTP scenario based on an entrance fee, which can often be problematic (and possibly send out politically unhelpful messages) in a cultural setting. The hybrid design employed – where compensation should only *just* compensate respondents for an impact on wellbeing and crucially only for those that said their wellbeing would be negatively affected in the first place – seemingly has the added advantage **of reducing the well-known disparities between WTA and WTP valuations** that are often found in the literature (something we also found, though to a lesser extent, in the NHM study). As such, the Tate Liverpool study adds to the evidence from our NHM study that WTP and hybrid versions of the WTA questions perform well when used to provide economic valuations of the work of cultural institutions.

Although WTP has now become the preferred monetary elicitation method in the CV literature, it is acknowledged that there are times when WTA is warranted, for example when respondents come from very poor backgrounds, say, such that their WTP amounts are constrained, or when property rights are such that respondents believe they have some *intrinsic* right to the good or service in question (culture is arguably just such a case). **In instances like these we have shown that a hybrid approach grounded also in the theory of WV can deliver plausible values.**

## 12. Overall conclusions

In the concluding section we discuss our findings in relation to their potential policy implications and to the research questions presented at the outset of the paper.

### 12.1. Interpretation of CV results

The final estimates presented in Tables 7.1 and 11.1 relate to **three types of economic value for the work done respectively by the Natural History Museum and Tate Liverpool**. It is important to reiterate what these values refer to, and how they can be used in discussions of value in culture.

First, following best-practice in CV methodology, our case studies show that the method is a **viable approach for measuring economic values in the context of cultural institutions**. The method produced **realistic values** that varied in ways predicted by theory and by previous findings, **across different institutions, scenarios, payment vehicles and population groups**.

Second, the values **are specific to the cultural institution from which they are derived**. In this case, the NHM (Table 7.1) a large internationally important museum of natural history in the capital city of the UK, and Tate Liverpool (Table 11.1), a prominent art gallery in a major city in the Northwest of England. Crucially, these values are not **transferable between these two institutions and they should therefore not be compared**. Not only are there considerable differences between the institutions themselves, and their work inside and outside their buildings, but also the populations surveyed are different (the institutions have different types of visitors as depicted in Table 12.1, with TL visitors being generally older than NHM visitors, more likely to be female, having higher household income, fewer dependent children, and being more likely to be employed, and a member of a heritage, conservation, or environmental organisation; and the general populations surveyed in each case also have different regional weightings). There are also significant differences in the scenarios employed in the contingent valuations (the payment mechanisms are different between case studies for example, with entry fees being used in the NHM visitor survey and donations being used in the TL visitor survey). Value transfer is only valid in cases where the institutions, the changes to be valued, and the populations affected are similar, which is not the case for the two institutions considered here.

The concomitant is that there is, however, potential to use the TL values estimated here to predict the values assigned to similar regional art galleries, facing similar changes. Or to use the estimated NHM values to predict the values of similarly large international natural history museums, facing similar changes. For a thorough empirical study of value transfer in cultural settings, we refer readers to the recent EU project Climate for Culture ([www.climateforculture.eu](http://www.climateforculture.eu)).

**Table 12.1 Comparison of respondents in the NHM and TL visitor surveys**

	NHM visitor survey (weighted)	TL visitor survey (weighted)
Male (%)	51	43
Age (mean)	39	47
Household income (£, mean)	39,000	41,000
Dependent children under 16 years (%)	35	15
Married/with partner (%)	40	43
University education (%)	62	68
In employment (full-time, part-time, self-employed) (%)	47	65
Health (good, very good, excellent)(%)	94	92
Member (heritage, conservation, environmental org) (%)	17	37
Total	616	526

Lastly, **care must be taken when using and interpreting the range of values estimated.** For each institution we have estimated three core types of WTP value (Tables 7.1 and 11.1), with different meanings depending on the context they refer to, and varying further within sub-samples:

- A. Direct use value for current visitors (as an entry fee in the case of the NHM and as a donation in the case of TL)
- B. Non-use value for current visitors (as a voluntary top up to an entry fee to support research and conservation work in the case of the NHM; and as a separate donation to support outreach community work in the case of TL)
- C. Non-use and option values for the general UK population (excluding N. Ireland) as an annual donation, for the work of each of the institutions, inside and outside their buildings.

We estimate additionally WTA values to compensate for a hypothetical closure of each institution (Tables 6.23 and 10.22).

**These values are all conceptually distinct, referring to different scenarios, payment settings, and population groups. It is therefore important that each value is used in the correct policy context.**

## 12.2 Can preference-based and wellbeing valuation methods deliver convincing valuations in cultural institutions?

In this research project, we have found that contingent valuation – perhaps the most commonly used form of preference-based valuation – performs well in measuring use and non-use/option values for both Tate Liverpool and the NHM, when undertaken using best-practice methodology to address its well-known implementation challenges. In particular, our surveys have generated plausible and theoretically consistent estimates of WTP for access to the cultural institutions (using either entry fees or donations) and for donations for non-use elements of the institution, such as community work in the case of Tate Liverpool and conservation and research in the case of the NHM.

We employ the following best practice elements which we suggest should be seen as the standard requirement for CV as applied in the cultural sector:

- Detailed **descriptions of the good/service** including **photos**.
- Detailed **description of the payment vehicle**.
- Using **survey pre-testing** to check familiarity with the survey information and any difficulties in answering questions.
- Use of **the most appropriate payment mechanisms to the circumstances of each site** (for example entry fees, top up donations on top of entry fees, or stand-alone voluntary donations) to elicit direct use values and/or non-use values.
- Use of a **wide payment range in the payment ladder** (informed by the results of survey testing) to reduce anchoring and range effects. We note that UK guidelines (Bateman et al., 2002) state that **single bounded dichotomous choice** is preferable for large samples. But **payment cards** are also included in UK guidelines, and provide a user-friendly alternative method for eliciting WTP values for cultural institutions that work well with smaller samples.
- Use of **cheap talk and oath scripts** to reduce hypothetical bias.
- **Consider using WTA when WTP is problematic, adopting the hybrid contingent-wellbeing valuation approach suggested here.**

We have made an important caveat to our findings that WTP for access may have been primed by the entrance fees to paid exhibitions and/or the gallery's donation suggestions, and in this sense the plausibility of the results from the CV study may not be so surprising after all. A question remains therefore whether CV – even with the best practice considerations set out here – would derive equally plausible or convincing values in cultural institutions in instances where there is no relevant market price that could act as a feasible benchmark (such as when they house unique and iconic collections). This is an important area for future research.

The **choice of payment mechanism** is context-dependent, and subject to practical constraints. For the purposes of capturing visitor use values, an **entry fee is the ideal payment mechanism as it avoids the known problems of voluntary payments** (Bateman et al., 2002). However, this may not be applicable in many cases, as where the cultural institution is sensitive to creating any impression in the minds of the public that entry fees may in the future be introduced. In these cases, and while the hypothetical nature of the contingent valuation exercise should always be stressed throughout the survey, a **donation payment mechanism** may be used as a **second best option (as is the case in the TL study)**. Moreover, the **hybrid contingent-wellbeing valuation approach suggested here, based on a WTA format, may also be used.**

We look at **three variants of the WV method**.

The standard WV method proves to be problematic in both the TL and NHM cases. For TL, we find that visits to the gallery are not statistically significant determinants of life satisfaction, whereas in the case of the NHM we find that the estimated impact of visits on life satisfaction is implausibly high, due possibly to focusing illusion effects (because people are being reminded about the NHM during the survey). This contrasts with previous studies in this area that have used large national datasets such as Taking Part and have found plausible values for cultural engagement using the WV method.

The problems that we have encountered may conceivably be due to a combination of smaller sample sizes and focusing illusion effects, issues that the national surveys do not suffer from. Equally, it may be that the life satisfaction impacts of engaging with a single cultural institution will always be too small to detect. Therefore, we conclude within the context of this study, that the standard approach to WV should not yet be applied to value the work of individual cultural institutions, even if WV studies using national data have been shown to give plausible results for the sector as a whole. It will be very important in future research, however, to test the robustness of this conclusion with other cultural organisations, especially with larger sample surveys than has been possible in the present study.

The vignettes approach to WV, which we apply to culture for the first time, produces decidedly mixed results, with promising results in the NHM case study, but no significant results in the case of TL. This is another area where further research is worthwhile.

**A striking finding from our research is that the hybrid contingent-wellbeing valuation approach, combining elements of both CV and WV – where compensation should only *just* compensate respondents for an impact on wellbeing and crucially only for those that said their wellbeing would be negatively affected in the first place – is shown to work well in both the NHM and TL case studies.** In particular, the well-known disparity in WTA and WTP estimates which plague pure CV studies are found to be reduced. The explicit reference to wellbeing – in the form of life satisfaction – seems to generate plausible values per visit to TL and the NHM. In practice, there are in any case many cultural settings where a WTA payment format is preferable to WTP, so a potentially important conclusion is that the hybrid valuation method we have tested is a promising approach.

### 12.3 Some further methodological considerations

An important methodological concern in both the NHM and TL surveys relates to the risk of overstatement in certain responses. For example, we find that a higher proportion of respondents place heritage and culture in their top three public spending priorities than we would perhaps expect. This is true in both the online and visitor surveys, and can be accounted for by the framing of the questionnaire instrument around the value of museums and galleries. Intuitively, an even higher proportion of respondents prioritise heritage and culture in the visitors' surveys. Similar results have been found in other recent heritage visitor surveys ([www.climateforculture.eu](http://www.climateforculture.eu)).

Another significant overstatement is found in the general population surveys, where respondents are asked if they have visited the institution in question in the previous 12 months. In both the NHM and TL studies, the proportion of the online general population sample that claim to have visited the cultural institution in the previous year is much higher than plausible, given annual visitor numbers to these institutions. A number of considerations should be borne in mind here. First, the online panels are **not representative of visitors** (i.e. we did not set a quota for visitation) and may therefore have oversampled them for some reason unaccounted for, such as the amount of spare time available to complete the survey. Second, we may have identified a **recollection bias**, whereby respondents struggle to recall events and incorrectly assign past visits to the last 12 months when in fact they have happened earlier. Third, there is the possibility that we are capturing a **'yea-saying' effect**, whereby respondents say that they have visited in the past 12 months because it seems simply the easiest answer to give. Confounding these possibilities, there may also have been financial incentive in the online sample to stay in the survey: respondents are paid a small fee by the survey company each time they complete a survey, and they may have (incorrectly) thought that answering no to this question would have led to them being filtered out of the survey.

In the case of Tate Liverpool, we attempt to minimise the likelihood of yea-saying by revealing the Tate Liverpool focus after the visit frequency question has been asked, and by asking follow-up questions to those who have indicated that they have visited in the previous 12 months, though this does not eradicate the problem. A notable finding in the Tate Liverpool study is that overstatement is especially high for respondents from the Northwest of England. This may plausibly reflect how the survey was framed when respondents were being recruited, and an association between the cultural institution under study and factors like civic pride, personal interest/use. While these biases cast doubt on the reliability of the general population responses to these particular questions, they should not affect the hypothetical exercises used to elicit WTP, which, as we have seen, perform well

in standard validity tests. We would recommend that future research considers these framing and ordering biases in greater detail, and test for improved phrasing and follow-up questions.

#### **12.4. How consistent or conflicting are the valuations derived from applying these two methods?**

Contingent valuation methods give plausible and theoretically consistent estimates of WTP for entry fee/donations in both the Natural History Museum and Tate Liverpool case studies. Willingness to pay amounts are driven by theoretically consistent factors across all surveys, and in line with real life donation and entry fee amounts.

The hybrid contingent-wellbeing valuation approach produces WTA per visit values for both the Natural History Museum and Tate Liverpool that are similar to the respective visitor entry fee and donation amounts according to the contingent valuation surveys. This gives further confidence in the contingent valuation results, while at the same time reducing the disparities between WTA and WTP valuations that are commonly found in the literature.

#### **12.5. Can techniques from the two valuation methods be combined to optimise their use by cultural institutions and funders?**

What we have termed the hybrid contingent-wellbeing valuation method offers one potential approach for combining the best of what contingent and wellbeing valuation methodologies have to offer to cultural institutions like museums and galleries. The hybrid method uses a WTA format rather than WTP, which may be especially preferable under circumstances where respondents have constrained levels of income and/or when property rights dictate that people (at least perceive) that they have a right to the cultural good or service in question.

WV takes perhaps a more explicit approach than CV by couching the valuation task in direct reference to changes in people's wellbeing. The hybrid model brings this theoretical approach into the CV question design by asking people to provide a value for a good/service to compensate for changes in their subjective wellbeing – for those that have indicated that their subjective wellbeing is affected. The suggestion is that this might remind (or encourage) people to act in accordance with economic theory, which could result in more plausible values for non-market goods. Moreover, restricting the compensation to those that previously revealed that their subjective wellbeing is affected reduces instances of excessive compensation claims ('free lunches').

Future studies should test different wording formats for the hybrid contingent-wellbeing valuation approach, including the viability of its use in WTP questions as well. Moreover, the hybrid approach could be based on happiness rather than life satisfaction. The hybrid approach could also be used in principle to calibrate other WTA values that have been found to be of orders of magnitude higher than WTP estimates, through a conversion factor based on the difference in WTA estimates found in the standard CV and the hybrid method.

### 13. References

- Aadland, D., Caplan, A.J., 2006. Cheap talk reconsidered: New evidence from CVM. *J. Econ. Behav. Organ.* 60, 562–578. doi:10.1016/j.jebo.2004.09.006
- Angelini, V., Cavapozzi, D., Corazzini, L., Paccagnella, O., 2012. Age, Health and Life Satisfaction Among Older Europeans. *Soc. Indic. Res.* 105, 293–308. doi:10.1007/s11205-011-9882-x
- Ariely, D., 2000. Controlling the Information Flow: Effects on Consumers' Decision Making and Preferences. *J. Consum. Res.* 27, 233–248. doi:10.1086/jcr.2000.27.issue-2
- Armbrecht, J., 2014. Use value of cultural experiences: A comparison of contingent valuation and travel cost. *Tour. Manag.* 42, 141–148. doi:10.1016/j.tourman.2013.11.010
- Arrow, K., Solow, R., 1993. Report of the NOAA Panel on Contingent Valuation. National Oceanic and Atmospheric Administration Washington, DC.
- Arts Council England, 2014. The Value of Arts and Culture to People and Society: An Evidence Review. AC England, Manchester, UK.
- Báez-Montenegro, A., Bedate, A.M., Herrero, L.C., Sanz, J.Á., 2012. Inhabitants' willingness to pay for cultural heritage: a case study in Valdivia, Chile, using contingent valuation. *J. Appl. Econ.* 15, 235–258. doi:10.1016/S1514-0326(12)60011-7
- Barton, D. N. (2002). The transferability of benefit transfer: contingent valuation of water quality improvements in Costa Rica. *Ecol. Econ.* 42, 147–164. doi:10.1016/S0921-8009(02)00044-7
- Bateman, I.J., Carson, R.T., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Özdemiroğlu, E., Pearce, D.W., Sugden, R., Swanson, J., 2002. *Economic Valuation with Stated Preference Techniques: A Manual.* 458.
- Bateman, I.J., Day, B.H., Jones, A.P., Jude, S., 2009. Reducing gain–loss asymmetry: A virtual reality choice experiment valuing land use change. *J. Environ. Econ. Manag.* 58, 106–118. doi:10.1016/j.jeem.2008.05.003
- Bedate, A., Herrero, L.C., Sanz, J.Á., 2004. Economic valuation of the cultural heritage: application to four case studies in Spain. *J. Cult. Herit.* 5, 101–111. doi:10.1016/j.culher.2003.04.002
- 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. *J. Cult. Econ.* 33, 185–199. doi:10.1007/s10824-009-9098-y
- Beja, E.J., 2011. Subjective well-being approach to the valuation of income inequality (MPRA Paper).
- Bertrand, M., Mullainathan, S., 2001. Do People Mean What They Say? Implications for Subjective Survey Data. *Am. Econ. Rev.* 91, 67–72.
- Blanchflower, D.G., Oswald, A.J., 2004. Well-being over time in Britain and the USA. *J. Public Econ.* 88, 1359–1386. doi:10.1016/S0047-2727(02)00168-8
- Bockstael, N.E., McConnell, K.E., 1980. Calculating Equivalent and Compensating Variation for Natural Resource Facilities. *Land Economics.* 56:56–63
- Boter, J., Rouwendal, J., Wedel, M., 2005. Employing Travel Time to Compare the Value of Competing Cultural Organizations. *J. Cult. Econ.* 29, 19–33. doi:10.1007/s10824-005-5796-2
- Brida, J.G., Meleddu, M., Pulina, M., 2012. Factors influencing the intention to revisit a cultural attraction: The case study of the Museum of Modern and Contemporary Art in Rovereto. *J. Cult. Herit.* 13, 167–174. doi:10.1016/j.culher.2011.08.003
- Brida, J.G., Monterubbianesi, P.D., Zapata-Aguirre, S., 2013. Determinants of visitor's satisfaction to a cultural attraction: the case of the museum of Antioquia, Colombia. *Estud. Perspect. En Tur.* 22, 729–744.
- Bridaa, J.G., Meleddub, M., Pulinac, M., 2012. Understanding Urban Tourism Attractiveness: The Case of the Archaeological Ötzi Museum in Bolzano. *J. Travel Res.* 0047287512437858. doi:10.1177/0047287512437858
- Brouwer, R. 2000. Environmental value transfer: State of the art and future prospects. *Ecol. Econ.* 32(1), 137-152.
- Burton, C., Louviere, J., Young, L., 2009. Retaining the visitor, enhancing the experience: identifying attributes of choice in repeat museum visitation. *Int. J. Nonprofit Volunt. Sect. Mark.* 14, 21–34. doi:10.1002/nvsm.351

- Cameron, T.A., Huppert, D.D., 1989. OLS versus ML estimation of non-market resource values with payment card interval data. *J. Environ. Econ. Manag.* 17, 230–246. doi:10.1016/0095-0696(89)90018-1
- Carlsson, F., Frykblom, P., Johan Lagerkvist, C., 2005. Using cheap talk as a test of validity in choice experiments. *Econ. Lett.* 89, 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. *J. Econ. Behav. Organ.* 89, 105–121. doi:10.1016/j.jebo.2013.02.003
- 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. *J. Econ. Behav. Organ.* 89, 105–121. doi:10.1016/j.jebo.2013.02.003
- 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. *J. Econ. Behav. Organ.* 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?
- Carroll, N., Frijters, P., Shields, M.A., 2009. Quantifying the costs of drought: new evidence from life satisfaction data. *J. Popul. Econ.* 22, 445–461. doi:10.1007/s00148-007-0174-3
- Carson, R.T., 2000. Contingent Valuation: A User's Guide. *Environ. Sci. Technol.* 34, 1413–1418. doi:10.1021/es990728j
- Carson, R.T., 2012. Contingent valuation: a practical alternative when prices aren't available. *J. Econ. Perspect.* 26, 27–42. doi:10.1257/jep.26.4.27
- Carson, R.T., Flores, N.E., Meade, N.F., 2001. Contingent valuation: controversies and evidence. *Environ. Resour. Econ.* 19, 173–210. doi:10.1023/A:1011128332243
- Champ, P.A., Bishop, R.C., 2001. Donation payment mechanisms and contingent valuation: an empirical study of hypothetical bias. *Environ. Resour. Econ.* 19, 383–402.
- Choi, A.S., Ritchie, B.W., Papandrea, F., Bennett, J., 2010. Economic valuation of cultural heritage sites: A choice modeling approach. *Tour. Manag.* 31, 213–220. doi:10.1016/j.tourman.2009.02.014
- Clark, A.E., Oswald, A.J., 2002. A simple statistical method for measuring how life events affect happiness. *Int. J. Epidemiol.* 31, 1139–1144. doi:10.1093/ije/31.6.1139
- Clawson, M., Knetsch, J.L., 2013. *Economics of Outdoor Recreation*. Routledge.
- Cohen, M.A., 2008. The Effect of Crime on Life Satisfaction. *J. Leg. Stud.* 37, S325–S353. doi:10.1086/597200
- Crossick, G., Kaszynska, P., 2014. Under construction: Towards a framework for cultural value. *Cult. Trends* 23, 120–131. doi:10.1080/09548963.2014.897453
- Cummings, R.G., Taylor, L.O., 1999. Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method. *Am. Econ. Rev.* 89, 649–665.
- Dasgupta, P., Serageldin, I., 1999. *Social Capital: A Multifaceted Perspective*. World Bank.
- De Rojas, C., Camarero, C., 2008. Visitors' experience, mood and satisfaction in a heritage context: Evidence from an interpretation center. *Tour. Manag.* 29, 525–537. doi:10.1016/j.tourman.2007.06.004
- de-Magistris, T., Pascucci, S., 2014. The effect of the solemn oath script in hypothetical choice experiment survey: A pilot study. *Econ. Lett.* 123, 252–255. doi:10.1016/j.econlet.2014.02.016
- Deaton, A., 2008. Income, health and wellbeing around the world: Evidence from the Gallup World Poll. *J. Econ. Perspect. J. Am. Econ. Assoc.* 22, 53.
- Del Chiappa, G., Ladu, M.G., Meleddu, M., Pulina, M., 2013. Investigating the degree of visitors' satisfaction at a museum. *Anatolia* 24, 52–62. doi:10.1080/13032917.2012.762317
- Del Saz Salazar, S., Montagud Marques, J., 2005. Valuing cultural heritage: the social benefits of restoring and old Arab tower. *J. Cult. Herit.* 6, 69–77. doi:10.1016/j.culher.2004.09.001
- Di Tella, R., MacCulloch, R., 2006. Some Uses of Happiness Data in Economics. *J. Econ. Perspect.* 20, 25–46.

- Diamond, P.A., Hausman, J.A., 1994. Contingent Valuation: Is Some Number better than No Number? *J. Econ. Perspect.* 8, 45–64.
- Diener, E.D., 1984. Subjective well-being. *Psychol. Bull.* 95, 542–575.
- Dolan, P., Fujiwara, D., 2012. Valuing adult learning: comparing wellbeing valuation to contingent valuation.
- Dolan, P., Layard, R., Metcalfe, R., 2011. Measuring subjective well-being for public policy.
- Dolan, P., Metcalfe, R., 2008. Comparing Willingness-to-pay and Subjective Well-being in the Context of Non-Market Goods, CEP Discussion Paper No 890
- Dolan, P., Peasgood, T., White, M., 2008. Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *J. Econ. Psychol.* 29, 94–122. doi:10.1016/j.joep.2007.09.001
- 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. Economics for the Environment Consultancy, London, UK.
- Ehmke, M. D., Lusk, J. L., & List, J. A., 2008. Is Hypothetical Bias a Universal Phenomenon? A Multinational Investigation. *Land Economics*, 84, 489-500.
- Ferreira, S., Moro, M., 2010. On the Use of Subjective Well-Being Data for Environmental Valuation. *Environ. Resour. Econ.* 46, 249–273. doi:10.1007/s10640-009-9339-8
- Environmental Change. <http://dx.doi.org/10.1016/j.gloenvcha.2013.03.010>
- Ferrer-i-Carbonell, A., Van Praag, B., 2002. The subjective costs of health losses due to chronic diseases. An alternative model for monetary appraisal. *Health Econ.* 11, 709–722.
- Fonseca, S., Rebelo, J., 2010. Economic valuation of cultural heritage: application to a museum located in the Alto Douro wine region–World Heritage Site. *PASOS Rev. Tur. Patrim. Cult.* 8, 339–350.
- Freeman, A.M., 2003. Economic valuation: What and why, in: *A Primer on Nonmarket Valuation*. Kluwer Academic Publishers, pp. 1–25.
- Frey, B.S., Luechinger, S., 2004. Decentralization as a disincentive for terror. *Eur. J. Polit. Econ.*, The Economic Consequences of Terror 20, 509–515. doi:10.1016/j.ejpoleco.2004.03.001
- Fujiwara, D., Dolan, P., forthcoming. Happiness-Based Policy Analysis. *Handbook of Well-being and Public Policy*. Oxford University Press.
- Fujiwara, D., & MacKerron, G. (2015). Cultural activities, artforms and wellbeing (p. 37). London, UK: Smetrica, Arts Council England. [http://media.wix.com/ugd/9ccf1d\\_05c61fa6c3ee4ea68de17ce62f71cd53.pdf](http://media.wix.com/ugd/9ccf1d_05c61fa6c3ee4ea68de17ce62f71cd53.pdf)
- Fujiwara, D., 2013. The social impact of housing providers. HACT, Legal & General, Plus Dane Group, London, UK.
- Fujiwara, D., 2013a. Museums and Happiness: The Value of Participating in Museums and the Arts. The Happy Museum Project, London, UK.
- Fujiwara, D., 2013b. A General Method for Valuing Non-Market Goods Using Wellbeing Data: Three-Stage Wellbeing Valuation, in: CEP Discussion Paper No 1233. Centre for Economic Performance, London School of Economics, London, UK, pp. 1–29.
- Fujiwara, D., Campbell, R., 2011. Valuation Techniques for Social Cost-Benefit Analysis: Stated Preference, Revealed Preference and Subjective Well-Being Approaches. A Discussion of the Current Issues. HM Treasury, London, UK.
- Fujiwara, D., Dolan, P., 2014. Valuing mental health. *Policy* 4, 2.1.
- Fujiwara, D., Kudrna, L., Dolan, P., 2014. Quantifying and Valuing the Wellbeing Impacts of Culture and Sport. Department for Culture, Media & Sport, London, UK.
- Gilbert, D., 2007. *Stumbling on Happiness*. Harper Perennial. London.
- Groot, W., van den Brink, H.M., 2006. The compensating income variation of cardiovascular disease. *Health Econ.* 15, 1143–1148. doi:10.1002/hec.1116
- Hanley, N., Mourato, S., Wright, R.E., 2001. Choice Modelling Approaches: A Superior Alternative for Environmental Evaluation?. *J. Econ. Surveys*. 15, 435-462.
- Hausman, J., 2012. Contingent valuation: from dubious to hopeless. *J. Econ. Perspect.* 26, 43–56. doi:10.1257/jep.26.4.43

- Hicks, J.R., Allen, R.G.D., 1934. A reconsideration of the theory of value. Part I. *Economica* 1, 52–76.
- Horowitz, J.K., McConnell, K.E., 2002. A Review of WTA/WTP Studies. *J. Environ. Econ. Manag.* 44, 426–447. doi:10.1006/jeem.2001.1215
- Hotelling, H., 1949. Letter quoted by R. E. Prewitt in “Economic Study of the Monetary Evaluation of Recreation in National Parks.” United States Department of Interior, Washington, D.C.
- Hsee, C.K., Zhang, J., 2004. Distinction Bias: Misprediction and Mischoice Due to Joint Evaluation. *J. Pers. Soc. Psychol.* 86, 680–695. doi:10.1037/0022-3514.86.5.680
- Jaffry, S., Apostolakis, A., 2011. Evaluating individual preferences for the British Museum. *J. Cult. Econ.* 35, 49–75. doi:10.1007/s10824-010-9133-z
- Jeong, J.-H., Lee, K.-H., 2006. The physical environment in museums and its effects on visitors’ satisfaction. *Build. Environ.* 41, 963–969. doi:10.1016/j.buildenv.2005.04.004
- Johns, H., Ormerod, P., 2007. Happiness, Economics and Public Policy (SSRN Scholarly Paper No. ID 1020246). Social Science Research Network, Rochester, NY.
- Jura Consultants, 2005. Bolton’s Museum, Library and Archive Services; An Economic Valuation. Lond. Mus. Libr. Arch. Council.
- Kahneman, D., 2000. Evaluation by moments: Past and future, in: Kahneman, D., Tversky, A. (Eds.), *Choices, Values, and Frames*. Cambridge University Press, Cambridge UK, pp. 693–708.
- Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., Stone, A. A. (2004). A survey method for characterizing daily life experience: the day reconstruction method. *Science*, 306 1776–1780. doi:10.1126/science.1103572
- Kahneman, D., Krueger, A.B., 2006. Developments in the measurement of subjective well-being. *J. Econ. Perspect.* 20, 3–24. doi:10.1257/089533006776526030
- Kahneman, D., Krueger, A.B., Schkade, D., Schwarz, N., Stone, A.A., 2006. Would You Be Happier If You Were Richer? A Focusing Illusion. *Science* 312, 1908–1910. doi:10.1126/science.1129688
- Kahneman, D., Snell, J., 1992. Predicting a changing taste: Do people know what they will like? *J. Behav. Decis. Mak.* 5, 187–200. doi:10.1002/bdm.3960050304
- Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47, 263–291. doi:10.2307/1914185
- Kapteyn, A., Smith, J.P., Soest, V., Arthur, Vonkova, H., 2011. Anchoring Vignettes and Response Consistency (SSRN Scholarly Paper No. ID 1799563). Social Science Research Network, Rochester, NY.
- Kim, S.S., Wong, K.K.F., Cho, M., 2007. Assessing the economic value of a world heritage site and willingness-to-pay determinants: A case of Changdeok Palace. *Tour. Manag.* 28, 317–322. doi:10.1016/j.tourman.2005.12.024
- King, D.M., Mazzotta, M.J., 2011. *Ecosystem Valuation: Contingent Valuation Method*.
- King, G., Murray, C. J. L., Salomon, J. A., Tandon, A. (2004). Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research. *American Political Science Review.* 127, 421–433. doi:10.1017/S000305540400108X
- Krosnick, J.A., 1991. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Appl. Cogn. Psychol.* 5, 213–236. doi:10.1002/acp.2350050305
- Lampi, E., Orth, M., 2009. Who visits the museums? A comparison between stated preferences and observed effects of entrance fees. *Kyklos* 62, 85–102. doi:10.1111/j.1467-6435.2009.00425.x
- Lichtenstein, S., Slovic, P., 2006. *The Construction of Preference*. Cambridge University Press, Cambridge UK.
- List, J.A., 2003. Does market experience eliminate market anomalies? *Q. J. Econ.-Camb. Mass.-* 118, 41–72.
- List, J.A., Lucking-Reiley, D., 2000. Demand reduction in multiunit auctions: Evidence from a sportscard field experiment. *Am. Econ. Rev.* 961–972.
- Loewenstein, G., Adler, D., 1995. A Bias in the Prediction of Tastes. *Econ. J.* 105, 929–937. doi:10.2307/2235159
- Loewenstein, G., Ubel, P.A., 2008. Hedonic adaptation and the role of decision and experience utility in public policy. *J. Public*

- Econ., Special Issue: Happiness and Public Economics 92, 1795–1810. doi:10.1016/j.jpube.2007.12.011
- Lusk, J.L., 2003. Effects of Cheap Talk on Consumer Willingness-to-Pay for Golden Rice. *Am. J. Agric. Econ.* 85, 840–856. doi:10.1111/1467-8276.00492
- MacKerron, G., Mourato, S. 2013. Happiness is greater in natural environments. *Global Environmental Change*. <http://dx.doi.org/10.1016/j.gloenvcha.2013.03.010>
- Maddison, D., Foster, T., 2003. Valuing congestion costs in the British Museum. *Oxf. Econ. Pap.* 55, 173–190. doi:10.1093/oep/55.1.173
- Maddison, D., Mourato, S., 2001. Valuing different road options for Stonehenge. *Conserv. Manag. Archaeol. Sites* 4, 203–212. doi:10.1179/135050301793138182
- Malpezzi, S., 2003. Hedonic pricing models: a selective and applied review. *Sect. Hous. Econ. Public Policy Essays Honor Duncan MacLennan*.
- Martin, F., 1994. Determining the size of museum subsidies. *J. Cult. Econ.* 18, 255–270. doi:10.1007/BF01079759
- Mazzanti, M., 2003. Discrete choice models and valuation experiments. *J. Econ. Stud.* 30, 584–604. doi:10.1108/01443580310504453
- McFadden, D., 1974. The measurement of urban travel demand. *J. Public Econ.* 3, 303–328. doi:10.1016/0047-2727(74)90003-6
- McFadden, D., 1977. *Modelling the Choice of Residential Location* (Cowles Foundation Discussion Paper No. 477). Yale University, New Haven, CT.
- Metcalfe, R., Powdthavee, N., Dolan, P., 2011. Destruction and Distress: Using a Quasi-Experiment to Show the Effects of the September 11 Attacks on Mental Well-Being in the United Kingdom\*. *Econ. J.* 121, F81–F103. doi:10.1111/j.1468-0297.2010.02416.x
- Mitchell, R.C., Carson, R.T., 1989. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Resources for the Future.
- Moore, S.C., 2006. The value of reducing fear: an analysis using the European Social Survey. *Appl. Econ.* 38, 115–117. doi:10.1080/00036840500368094
- Morrison, M., Brown, T.C., 2009. Testing the Effectiveness of Certainty Scales, Cheap Talk, and Dissonance-Minimization in Reducing Hypothetical Bias in Contingent Valuation Studies. *Environ. Resour. Econ.* 44, 307–326. doi:10.1007/s10640-009-9287-3
- Mourato, S., Kontoleon, A., Danchev, A., 2002. *Preserving cultural heritage in transition economies: A contingent valuation study of Bulgarian monasteries*. Edward Elgar.
- Mourato, S., Kontoleon, A., Danchev, A., 2002. *Preserving cultural heritage in transition economies: A contingent valuation study of Bulgarian monasteries*. Edward Elgar.
- Mourato, S., Mazzanti, M., 2002. Economic valuation of cultural heritage: evidence and prospects. In: *Getty Conservation Institute, Assessing the Value of Cultural Heritage*. Getty Conservation Institute, Los Angeles.
- Murphy, J.J., Stevens, T., Weatherhead, D., 2003. *An Empirical Study of Hypothetical Bias in Voluntary Contribution Contingent Valuation: Does Cheap Talk Matter?* Unpubl. Manuscr.
- Navrud, S., Ready, R.C., 2002. *Valuing Cultural Heritage: Applying Environmental Techniques To Historic Buildings, Monuments and Artifacts*, Edward Elgar.
- Noonan, D.S., 2003. Contingent Valuation and Cultural Resources: A Meta-Analytic Review of the Literature. *J. Cult. Econ.* 27, 159–176. doi:10.1023/A:1026371110799
- O'Brien, D. (2010). *Measuring the Value of Culture: A Report to the Department for Culture Media and Sport*. London, UK: Department for Culture, Media and Sport. [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/77933/measuring-the-value-culture-report.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/77933/measuring-the-value-culture-report.pdf)

- O'Donnell, G., Deaton, A., Durand, M., Halpern, D., Layard, R. 2014. Wellbeing and Policy. Legatum Institute.
- O'Brien, 2010. Measuring the value of culture: A report to the Department for Culture, Media and Sport.
- Organisation for Economic Co-operation and Development (OECD). (2006). *Cost-Benefit Analysis and the Environment: Recent Developments*. Paris, France.
- Oswald, A.J., Powdthavee, N., 2008. Death, Happiness, and the Calculation of Compensatory Damages. *J. Leg. Stud.* 37, S217–S251. doi:10.1086/597200
- Packer, J., Bond, N., 2010. Museums as restorative environments. *Curator Mus. J.* 53, 421–436. doi:10.1111/j.2151-6952.2010.00044.x
- Parfit, D., 1984. *Reasons and Persons*. Oxford University Press, Oxford UK.
- Paswan, A.K., Troy, L.C., 2004. Non-Profit Organization and Membership Motivation: An Exploration in the Museum Industry. *J. Mark. Theory Pract.* 12, 1–15.
- Pearce, D., Atkinson, G., Mourato, S., 2006. *Cost-Benefit Analysis and the Environment: Recent Developments*. OECD Publishing.
- Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y., Podsakoff, N.P., 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *J. Appl. Psychol.* 88, 879–903. doi:10.1037/0021-9010.88.5.879
- Powdthavee, N., 2008. Putting a price tag on friends, relatives, and neighbours: Using surveys of life satisfaction to value social relationships. *J. Socio-Econ.* 37, 1459–1480. doi:10.1016/j.socec.2007.04.004
- Powdthavee, N., van den Berg, B. 2011. Putting different price tags on the same health condition: Re-evaluating the well-being approach. *J. Health Econ.* 30(5), 1032-1043.
- Pung, C., Clarke, A., Patten, L., 2004. Measuring the economic impact of the British library. *New Rev. Acad. Librariansh.* 10, 79–102. doi:10.1080/13614530412331296826
- Raharjo, M., Gravitiani, E., 2012. The Economic Value of Sangiran Museum, Central Java, Indonesia Application of Travel Cost Method. *Int. J. Soc. Sci. Econ. Art* 2, 39–43.
- Reeves, M., 2002. *Measuring the Economic and Social Impact of the Arts: A Review*. Arts Council England, London, UK.
- Rehdanz, K., Maddison, D., 2008. Local environmental quality and life-satisfaction in Germany. *Ecol. Econ.* 64, 787–797. doi:10.1016/j.ecolecon.2007.04.016
- Rolfe, J., Windle, J., 2003. Valuing the Protection of Aboriginal Cultural Heritage Sites. *Econ. Rec.* 79, S85–S95. doi:10.1111/1475-4932.00094
- Rondeau, D., D. Schulze, W., Poe, G.L., 1999. Voluntary revelation of the demand for public goods using a provision point mechanism. *J. Public Econ.* 72, 455–470. doi:10.1016/S0047-2727(98)00104-2
- Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. *J. Polit. Econ.* 34–55.
- Rubin, D.B., 2005. Causal Inference Using Potential Outcomes. *Journal of the American Statistical Association* 100, 322–331. doi:10.1198/016214504000001880
- Ruijgrok, E.C.M., 2006. The three economic values of cultural heritage: a case study in the Netherlands. *J. Cult. Herit.* 7, 206–213. doi:10.1016/j.culher.2006.07.002
- Santagata, W., Signorello, G., 2000. Contingent Valuation of a Cultural Public Good and Policy Design: The Case of "Napoli Musei Aperti". *J. Cult. Econ.* 24, 181–204. doi:10.1023/A:1007642231963
- Sanz, J.Á., Herrero, L.C., Bedate, A.M., 2003. Contingent Valuation and Semiparametric Methods: A Case Study of the National Museum of Sculpture in Valladolid, Spain. *J. Cult. Econ.* 27, 241–257. doi:10.1023/A:1026353218280
- Schwarz, N., Strack, F., 1999. Reports of subjective well-being: Judgmental processes and their methodological implications, in: *Well-Being: Foundations of Hedonic Psychology*. Russell Sage Foundation, London, UK, pp. 61–84.
- Shogren, J.F., Shin, S.Y., Hayes, D.J., Kliebenstein, J.B., 1994. Resolving Differences in Willingness to Pay and Willingness to Accept. *Am. Econ. Rev.* 84, 255–270.

- Smith, V.K., 2000. Fifty Years of Contingent Valuation. Paper presented at the Kobe Conference on the Theory and Application of Environmental Valuation, January 2000.
- Snowball, J.D., 2005. Art for the Masses? Justification for the Public Support of the Arts in Developing Countries – Two Arts Festivals in South Africa. *J. Cult. Econ.* 29, 107–125. doi:10.1007/s10824-005-5064-5
- Stevens, T.H., Tabatabaei, M., Lass, D., 2013. Oaths and hypothetical bias. *J. Environ. Manage.* 127, 135–141. doi:10.1016/j.jenvman.2013.04.038
- Stevenson, D., 2013. Reaching a “legitimate” value? A contingent valuation study of the National Galleries of Scotland. *Mus. Manag. Curatorship* 28, 377–393. doi:10.1080/09647775.2013.831250
- Stutzer, A., Frey, B.S., 2004. Reported subjective well-being: A challenge for economic theory and economic policy. *Schmollers Jahrb.* 124, 1–41.
- Sugden, R., 2005. Anomalies and Stated Preference Techniques: A Framework for a Discussion of Coping Strategies. *Environ. Resour. Econ.* 32, 1–12. doi:10.1007/s10640-005-6025-3
- Throsby, D., 1999. Cultural Capital. *J. Cult. Econ.* 23, 3–12. doi:10.1023/A:1007543313370
- Throsby, D., 2001. *Economics and Culture*. Cambridge University Press, Cambridge UK.
- Tohmo, T., 2004. Economic value of a local museum: Factors of willingness-to-pay. *J. Socio-Econ.* 33, 229–240. doi:10.1016/j.socec.2003.12.012
- Tuan, T.H., Navrud, S., 2008. Capturing the benefits of preserving cultural heritage. *J. Cult. Herit.* 9, 326–337. doi:10.1016/j.culher.2008.05.001
- Tusa, J., 2000. *Art matters: Reflecting on culture*. Methuen Pub Ltd.
- US Office of Management and Budget. (1992). Circular No. A-94 Revised (Transmittal Memo No. 64). Washington DC. [https://www.whitehouse.gov/omb/circulars\\_a094/](https://www.whitehouse.gov/omb/circulars_a094/)
- Van den Berg, B., Ferrer-i-Carbonell, A., 2007. Monetary valuation of informal care: the well-being valuation method. *Health Econ.* 16, 1227–1244. doi:10.1002/hec.1224
- Van Praag, B.M., Frijters, P., Ferrer-i-Carbonell, A., 2003. The anatomy of subjective well-being. *J. Econ. Behav. Organ.* 51, 29–49. doi:10.1016/S0167-2681(02)00140-3
- Veal, W.R., 2002. Content Specific Vignettes as Tools for Research and Teaching. *Electron. J. Sci. Educ.* 6.
- Welsch, H., 2002. Preferences over Prosperity and Pollution: Environmental Valuation based on Happiness Surveys. *Kyklos* 55, 473–494. doi:10.1111/1467-6435.00198
- Welsch, H., 2007. Environmental welfare analysis: A life satisfaction approach. *Ecol. Econ.* 62, 544–551. doi:10.1016/j.ecolecon.2006.07.017
- Welsch, H., 2008. The Social Costs of Civil Conflict: Evidence from Surveys of Happiness. *Kyklos* 61, 320–340. doi:10.1111/j.1467-6435.2008.00404.x
- Welsch, H., Kühling, J., 2009. Using happiness data for environmental valuation: issues and applications. *J. Econ. Surv.* 23, 385–406. doi:10.1111/j.1467-6419.2008.00566.x

## **Acknowledgements**

We would like to thank Mike Dixon and Justin Morris at the Natural History Museum, Andrea Nixon, Jemima Pyne and Sam Jones at Tate and Dan Culshaw at Morris Hargreaves McIntyre for their cooperation throughout this research. We are also grateful for the guidance and comments provided by Geoff Crossick, Patrycja Kaszynska, Mark Llewellyn, Paul Crawford, Joseph Lowe, Harman Sagger and Lee Miller. Thanks also go to John Davies at Nesta for his comments on an earlier draft of this report.

**The Cultural Value Project** seeks to make a major contribution to how we think about the value of arts and culture to individuals and to society. The project will establish a framework that will advance the way in which we talk about the value of cultural engagement and the methods by which we evaluate it. The framework will, on the one hand, be an examination of the cultural experience itself, its impact on individuals and its benefit to society; and on the other, articulate a set of evaluative approaches and methodologies appropriate to the different ways in which cultural value is manifested. This means that qualitative methodologies and case studies will sit alongside quantitative approaches.



Published by Arts and Humanities Research Council  
Polaris House, North Star Avenue, Swindon, Wiltshire, SN2 1FL  
[www.ahrc.ac.uk](http://www.ahrc.ac.uk)

©Arts and Humanities Research Council 2015  
Published October 2015

Design by Rumba.