

**Investigating the issue of hypothetical bias in Contingent Valuation studies.
Evidence from built heritage evaluation, in favor of CVM's validity**

S. Giannakopoulou¹, D. Damigos², D. Kaliampakos²

¹Aristotle University of Thessaloniki

²National Technical University of Athens

Abstract

Contingent Valuation (CV) method is the most widely used technique for the estimation of cultural goods due to its ability to capture their non-market value. However, the method's reliability receives, usually, criticism for a variety of biases, among which, the hypothetical bias remains the most controversial one. As relevant literature shows, respondents report higher WTP than they would actually pay. The paper aims at adding knowledge to the ongoing issue of hypothetical bias of CV studies in further support of the method's validity on the evaluation of cultural goods. In view of examining the core nature of hypothetical bias we conducted a meta-analysis of several CV surveys that were conducted in Greece, regarding the evaluation of the built heritage of traditional settlements. The results of the meta-analysis reveal that respondents give reasonable and repeatable answers throughout the questionnaires, even at the most sensitive questions. Findings that show this derive from answers regarding income, WTP as percentage of income, WTP in comparison to respondents' views on the preservation of built heritage, etc. Several findings of the meta-analysis show consistency of certain results, in the CV surveys, with the economic theory, i.e. the WTP amount in comparison to the offered quality of the cultural good. In addition, WTP estimations proved to be sensitive to a variety of external parameters, i.e. economic crisis. We argue, in this paper, that revealed and documented reliability throughout the surveys is not lost when it comes to the WTP amount statement; the core question of the CV survey. Therefore, although hypothetical bias remains, always, a controversial issue CVM is able to give, to a great extent, realistic results. Hence, in view of the elimination of hypothetical bias, questionnaires' structure should enable for testing respondents' reliability throughout the survey.

Keywords: Hypothetical Bias, Contingent Valuation, Willingness to pay, Cultural goods, Built heritage.

1. Introduction

Contingent Valuation (CV) is an accepted and well established monetary valuation method for non-market goods, widely applied in surveys in many countries and by many international organizations, i.e. World Bank, FAO, UNEP, OECD, since its introduction in 1963. Though it was initially targeted at environmental goods it is widely applied to cultural goods, mainly during the last twenty years (Bedate et al., 2012; Hoyos et al., 2010; Rizzo et al. 2006, Damigos et al., 2006; Venkatachalam, 2004; Noonan, 2003; Throsby 2000; Santagata et al., 2000). Since the only available methods for estimating passive-use values are the stated preference ones, their efficacy and credibility are repeatedly being put under questioning (Kling et al., 2012). CV studies receive criticism for a variety of issues concerning the reliability and credibility of goods' valuation. CV studies are prone to several biases, questioning their convenience in policy decision planning. Among the main biases, CV method is being constantly criticized on holding hypothetical bias thus, divergence between hypothetical and actual WTP values (Murphy et al., 2005; List and Gullet, 2001; Cummings et al., 1986).

Though documented progress in relative studies strengthens the method's credibility, hypothetical bias still remains at the core of controversy. The phrase "ask a hypothetical question to receive a hypothetical answer" precisely summarizes the problem. In other words, since respondents are not actually obliged to pay for the good provided they may give a false statement, either overestimating or underestimating their true WTP. Overestimation of WTP value is what mostly takes place (Arrow et al., 2003). For a variety of reasons people tend to state higher amounts of money than they are actually willing to give. Hence, estimated economic values lack veridicality. As it will be shortly presented later, a variety of methods have been developed aiming at hypothetical bias detection and elimination. Yet, there is ongoing research on the issue since a firm methodology effectively eliminating hypothetical bias still hasn't been established.

This paper focuses on analyzing a different approach testing hypothetical bias in CV studies. We argue that elaboration on a CV study key-findings, can effectively detect reliability in respondents' overall attitude to the survey. If overall credibility is proved to derive then isolating WTP statement and yielding it non true opposes rational argumentation. Thus, we argue that since "truth detection" in individual's WTP is not feasible, a methodology should be developed in order to capture statement credibility of responses that can be firmly tested. The paper presents a meta-analysis on the results of five CV studies applied in the evaluation of traditionally built environment. In specific, CV method was used in order to estimate the economic benefits of preserving local vernacular architecture in two mountainous settlements, in Greece. The aim of the studies was to reveal people's attitudes towards the good's protection and whether they would justify annual

investments on preservation works. Comparison and analysis of the five CV studies results led to the development of a methodology on testing respondents' statements validity.

The rest of this paper is structured as follows. In section 2, relative literature on hypothetical bias issue in CV studies is described. In section 3, the five CV studies framework and design structure are presented. Section 4 is consisted of the CV studies key-results used to the meta-analysis study, the meta-analysis study's methodology and its results. Section 5 concludes this paper pointing out elements for future research in the hypothetical bias field.

2. The issue of hypothetical bias in CV studies; relative literature

Important criticism is made on CV's reliability referring to several forms of implicit biases (Harrison et al., 2008; Ajzen et al., 2003; Carson et al., 2001; Bateman et al., 1997; Mitchell et al., 1989). The NOAA panel report set several rules in order to strengthen the CV's efficacy, when the method's reliability was put under the microscope after the Exxon Valdez oil spill (Kling et al., 2012, Arrow et al., 1993). Among the most controversial issues and weakest features of CV, the hypothetical bias prevails.

CV studies are based on the creation of a hypothetical market around the good in question in order to capture individual's willingness to pay (or accept) money for the provision (or the loss) of a good (public, quasi-public or private) (Tentes et al., 2012; Hoyos et al., 2010; Santagata et al., 2000; Hanemann, 1994; Mitchell et al., 1989). Since respondents to CV studies are not actually asked to pay the amount of money they state, the core problem rests in the effective *truth detection* in their hypothetical statements. Proof of such certainty would armor the method and significantly increase its credibility.

Existence of hypothetical bias and ways of establishing a firm methodology on its elimination keeps emerging in current literature. As relative studies show, people tend to overestimate their WTP values for a variety of reasons; willing to increase the chances of the good's provision, wishing to please the researcher, misunderstanding of the good valued or the presented hypothetical scenario etc. (Mohammed, 2012; Harrison et al., 2008). Comparisons between hypothetical and actual payment for the same good document that, in the majority of studies, hypothetical values override actual ones (Mohammed, 2012; Champ et al., 2001; List et al., 2001). Morrison et al. (2009) describe studies comparing hypothetical and actual WTP values showing ratios of 3.26 and 3.13. Carson et al. (1996) examined 83 studies with 616 comparisons between hypothetical and actual payments. They found out that the ratio between SP (Stated Preference) and RP (Revealed Preference) estimates lies near 1.00 and the estimates are highly correlated (Azevedo et al., 2000). Murphy et al. (2005) note that individuals are widely believed to overestimate their economic valuation by a factor

of two or three, yet their meta-analysis on hypothetical bias of 28 stated preference valuation studies results in lower factors (1.35 median ratio of hypothetical to actual value).

In view of hypothetical bias elimination methodology, several approaches have been developed. Cummings and Taylor, firstly introduced the “cheap talk script” technique, in 1999, in environmental goods. “Cheap talk script” gained popularity among researchers and was repeatedly applied in various alternatives (Mohammed, 2012; Champ et al., 2004; Brown et al., 2003; List, 2001). Blackburn et al. (1994) developed a statistical bias function in order to estimate the extension of value overstatement using socioeconomic characteristics. Carlsson et al. (1999) mention that, stated preference methods, CV among them, are generally validated with internal and external tests. External tests are based on comparisons between hypothetical and actual WTP values, while internal tests are based on the examination of transitivity, stability of preferences and monotonicity (insensitivity to scope). Brown et al. (2003) suggest asking respondents to carefully consider their budget constraint before they state their hypothetical bid or instructing them to state whether or not they will actually give the amount of money they state. Vossler et al. (2003) note that the research on comparison between hypothetical response and actual behavior could be summarized in four groups: (a) Comparisons of CV with indirect valuation methods based on observed behavior; (b) comparisons of CV responses with actual market transactions; (c) simulated market studies that compare one group presented with a hypothetical decision to another group presented with an analogous decision where money is actually exchanged; and (d) comparisons of hypothetical and actual voting behavior. Certainty calibration techniques are suggested and compared by Samnaliev et al. (2006), contributing to the discussion on the motivation behind uncertainty of respondents. Morrison et al. (2009) suggest the use of follow-up certainty scales, which can either be numerical or qualitative ones (e.g. respondents are either asked if they are probably / definitely sure or they are asked to define their certainty with a 1-10 point scale), the presentation of response categories beyond the simple *yes* and *no* and the use of “cheap talk script”, as well. Mohammed (2012) categorizes mitigation techniques in *ex-ante* (cheap talk) and *ex-post* (uncertainty adjustment). Loomis (2011) also identifies two general groups of methods, i.e. *ex-ante* and *ex-post* approaches. *Ex-ante* approaches focus on reducing hypothetical bias in the survey design, while *ex-post* on recoding or calibrating WTP responses to correct or reduce the stated WTP for hypothetical bias. *Ex-ante* approaches are grouped in three techniques: (a) construction of a market with a potential effect on respondent; (b) use of the “cheap talk”; and (c) telling the respondent that the good will be provided based on the results of the survey (i.e. provision point mechanisms). *Ex-post* approaches consist of certainty scales and calibration factors. In a more recent study (Carlsson et al., 2013) the “oath script” was introduced, in response to the “cheap talk script” technique, examining the effects on hypothetical bias by asking

respondents to swear or promise to answer truthfully, mimicking the act of taking an oath in a courtroom. Finally, Xie et al. (2013) summarize hypothetical bias elimination techniques in four groups: (a) the use of “cheap talk”; (b) the addition of questions in the hypothetical experiment in order to get more information on consumers’ real attitudes; (c) the combination of hypothetical with non-hypothetical experiments in order to obtain calibration factors of the hypothetical bias; and (d) the use of non-hypothetical experiments, in order to avoid hypothetical bias.

Research on hypothetical bias sheds light to an issue of high importance. Yet, a common accepted method eliminating hypothetical bias in CV studies has not been established. Azjen et al. (2004) note, that an individual’s intention is the closest predictor of final behavior. Murphy et al. (2004, 2005) underline the need for future research to elaborate on the causes of hypothetical bias and the need for a comprehensive theory on hypothetical bias to be developed. Loomis (2011) points out that what is still not understood are the causes of hypothetical bias and, establishing a widely accepted general theory on them represents core issue of the stated preferences economic analysis.

3. CVM application on five field studies

3.1 The site areas of the CV studies

This study is based on a meta-analysis review on five CV studies concerning hypothetical WTP for the protection of vernacular architecture in two mountainous settlements. The five successive CV field studies took place from 2008 to 2011, in the town of Metsovo and the village of Sirako, both located in the mountain range of Pindus, in Epirus, Greece.

The town of Metsovo is located at an altitude of 1.150m, almost 50km from the city of Ioannina. It is a town of about 3.500 permanent habitants (EL.STAT, 2011) depending its development on a variety of economic activities; tourism being the most robust one. It is estimated that, on average, 200.000 people visit Metsovo at an annual base¹. Metsovo holds a continuous, uninterrupted historical presence in the area, since 15thct. It was never abandoned and has met a rather high development rate during the last decades. It used to be of an exceptionally high quality building environment, in the past, which bears severe decline during the last years. A variety of reasons hold responsibility for this; inadequate building legislation, neglect and indifference for local architecture’s protection being the most affective of them. Yet, an important amount of vernacular buildings do exist and appropriate

¹Data from the Municipality of Metsovo and from local tourist agents.

restoration works could lead to efficient retrieve and enhancement of lost traditional characteristics (Giannakopoulou, 2012).

The village of Sirako is also located at an altitude of 1.150m in the range of Tzoumerka (in mountain Pindos), 50km from the city of Ioannina. Sirako is a small settlement of about 200 family houses² used as vacation residents, during the year. Its habitants live in nearby urban centers staying at Sirako for some time (weeks of months) annually. Sirako counts only a few years of tourism development based on small sized hostels housed in old, renovated buildings, run by local families. It is estimated that almost 10.000 visitors arrive in Sirako, annually³. Unlike Metsovo, Sirako was abandoned in late 1970s and did not gained its former robust social structure ever since. Yet, its habitants never really abandoned it; during the 1990s most of them returned in order to rebuild their old houses. Due to its way of development through the years, as well as to a variety of legislative and subsidization measures, Sirako not only did not lose its local architectural identity but represents, today, one of the best well-preserved traditional settlements in the country (Giannakopoulou, 2012).

3.2 The CV surveys' design

The CV method was applied in order to estimate economic benefits deriving from vernacular architecture's preservation, in both settlements. To this direction an investment proposal was set up in order for the individual's WTP to be estimated. Individuals were addressed with the hypothetical scenario of the foundation of an independent institution that would undertake all necessary restoration works (in each settlement). Respondents were asked to state their maximum WTP in order for the plan to be realized. The aggregated WTP estimates depend on family / group WTP and the total number of benefited households / groups.

The economic benefits deriving from the provision of the good are primarily received by local residents and visitors. Residents have the opportunity to enjoy the traditional settlement in their everyday life. Visitors travel and dedicate time and money in order to enjoy it. Therefore, they represented the population target group of the five surveys. More specifically, there were three CV studies addressed to visitors and two CV studies addressed to residents (Table 1). The five surveys took place with face-to-face interviews during which respondents were addressed with a questionnaire. Residents were mainly interviewed by a door-to-door survey. Visitors, present at the settlements at the time of the surveys, were interviewed during their sightseeing walking –cafes, squares, local shops, sites of attraction etc.⁴.

²The habitants of Sirako are 300 families (source: the Municipality of Sirako).

³Data from the Municipality of Sirako and local tourism agents.

⁴ Parts of the results of the surveys –surveys 1, 2 and 4- were presented at the 17th ACEI International Conference, in 2012, in Kyoto, Japan, at the Heritage Conference, in 2012, in Porto, Portugal and

Overall, the five surveys took place in a four year period, with a total 1.346 face-to-face interviews (Table 1). Respondents were randomly chosen in order for representative sample of each group to be ensured.

3.3 The CV studies' questionnaire

Two different types of questionnaires were used; type A addressed to residents and type B addressed to visitors. Further differentiation among each type did exist in order to capture corresponding differences between the two settlements. Thus, there were two versions of questionnaire of type A and two versions of type B. Nevertheless, a basic set of common questions was included in each questionnaire. Each questionnaire was subdivided into four sections. The first section referred to introductory questions on the subject and the goal of the survey. The second one referred on the views and preferences of the respondents to a variety of issues concerning the good valued. The third section consisted of the valuation scenario, the core of the CVM application, the WTP elicitation question and questions examining the reasons for the stated WTP. The last section gathered demographic data from the respondents.

All respondents were presented the same hypothetical scenario. Visitors and residents were informed on the current situation of the good and the monetary investments needed for its protection. The scenario was then addressed to them as follows: *“As you are aware of, the protection and preservation of vernacular architecture of the settlement demands a certain amount of building works, i.e. construction, repairing, restoration etc. at a constant basis. If an independent foundation was to be established with the obligation to undertake all this necessary work, would you be willing to voluntarily contribute financially in order to support this foundation?”*. The respondents could answer stating “Yes” or “No” and, *if yes*, they were additionally asked to state the amount of money they would be willing to offer.

Voluntary contribution was the payment vehicle used in all of the five surveys. Two elicitation techniques were used. The open-ended format was used in four of the surveys and the payment card was used in one survey (Table 1). PC format survey chronically followed the OE one and therefore, lower and upper WTP values in PC were chosen based on the previous OE survey.

Table 1. Summarized data from the five CV surveys

	SURVEYS				
	1st	2nd	3rd	4th	5th
Settlement	Metsovo	Metsovo	Sirako	Metsovo	Sirako
Year of survey	2008	2009	2009 - 2010	2010	2010 - 2011
Population addressed	Visitors	Residents	Residents	Visitors	Visitors
Population of interest	300.000 individuals	1.000 (households)	300 (households)	150.000 individuals	10.000 individuals
Number of questionnaires	305	260	175	366	240
ETP elicitation technique	OE	OE	OE	PC	OE

4. Methodology and results.

4.1. Methodology

The overall aim of the meta-analysis study is to prove evidence of respondents' credibility and consistency of attitude to the body of the CV surveys that leads to safe assumptions on hypothetical bias elimination. To this direction, several findings of the five surveys as well as a variety of respondents' socioeconomic data collected were viewed as key-elements for further examination.

A first element that played decisive role to the overall results was the fact that, in 2009, Greece entered a phase of severe economic crisis, during which, unemployment rates extended previous upper limits, drastic income reductions took place and increasing numbers of people struggle to survive living under poverty limits. Such general socioeconomic situation is expected to affect individuals' choices, especially those referring to economic issues, such as their willingness to donate money for the protection of heritage. A CV study examining people's financial choices should be able to capture and reflect the general social frame. At the beginning of this research only four CV surveys were designed to be applied. However, the sudden social change in 2009 led to the design and application of the fifth survey. Since the first CV survey took place in 2008, in the town of Metsovo targeted to its visitors, the fifth survey was decided to examine the same good and the same population group (i.e. visitors of Metsovo), in 2010. Two years after the first one and one year already with the impacts of economic crisis present.

A second element important for the purpose of this meta-analysis was the opportunity to examine the same cultural good (architectural heritage preservation) with five successive surveys, in two different case study areas. The opportunity to examine the same cultural good in five different surveys, during four successive years, gives the surveys the advantage of a repeated experiment (though not test-retest). This enables for more robust results to be obtained. Repeatability of findings shows solidity in social views, while time difference between surveys enables views' evolution detection.

Among several data collected from field survey a first finding considered important was the drastic reduction in the annual number of Metsovo's visitors, from 2008 to 2010. According to local authorities and hotel-managers⁵, in 2008 the estimated number of visitors in Metsovo was almost 300.000, while in 2010 was 150.000 (in Sirako, the annual estimated number of visitors at the time of the survey was estimated to 10.000 persons). It came out that in just two years of economic decline, there was a reduction of annual visitors to half. A reduction to the annual number of Greek visitors was indeed expected and logically explained; adopting a rather "safe mode" behavior in order to confront sudden income reductions and probable future shortcomings a cutback on luxury expenses, i.e. traveling, is a predictable attitude.

Several variables were considered as indicative and were further analyzed: the annual income variable, the occupational status variable, the WTP value variable and the (statistically important) variables affecting WTP. In addition, elicitation technique was also taken under consideration given that it affects the stated WTP value (Pan et al., 2012; Hoyos et al., 2010; Venkatachalam, 2004). Another variable examined was the "importance of the protection of the good", since it represents a basic motive for a positive WTP. The core working hypothesis of the study was that if respondents' statements document credibility and firmness in every testable answer then they should be attributed with high level of veridicality on their WTP value statement as well. Other than that, opposes rational argumentation. Hence, WTP value veridicality eliminates or may even exclude hypothetical bias. The study's examination axis was based on several testing comparisons and analysis of different variables. Methodologically, the study was based on three axes: (a) economic related variables and demographic data were analyzed in order for respondents' credibility to be effectively tested; (b) different questions, at different parts of the surveys were used as validity testing vehicles. Answers were cross-checked and respondents' reliability was tested; and (c) WTP value statements were tested according to economic theory frame. Results of the meta-analysis confirm the initial working hypothesis and establish a different methodological approach in detecting –and eliminating- hypothetical bias in CV studies.

⁵ Municipality of Metsovo, Community of Sirako and National Statistical Service of Greece.

4.2 Results of the meta-analysis study

The results of the meta-analysis study are organized according to different variables examined and their WTP value effect. Validity, inner coherence and economic theory consistency are tested. At a separate unit, WTP value is examined in relation to the offered quality of the good under the frame of economic theory principles.

4.2.1 The income factor

One issue of controversy, concerning CV surveys, is whether respondents answer truthfully questions referring to financial elements, in general (not only WTP values). The average annual income stated in each survey is shown in Table 2. Two findings concerning annual income were examined:

a. Comparing the average incomes between surveys 1 and 4 becomes evident that from 2008 to 2010 there was a mean reduction of annual stated average income of 14,5% (Table 2). This finding was then compared to relative data collected from statistical surveys held in Greece at that time and relative articles published on newspapers. According to that data, a mean income reduction of 15 – 16% was documented, at national level, due to wages' cutbacks. The surveys' findings prove that respondents' income statements hold credibility.

b. Annual income variable was subdivided into five categories: (a) Income < 10,000€, (b) 10,000€ < Income < 20,000€, (c) 20,000€ < Income < 30,000€, (d) 30,000 < Income < 40,000€, (e) Income > 40,000€. Comparisons between categories showed that there were statistically important differences at the 1st and 5th categories between Metsovo's visitors in 2008 and 2010. More specifically, more visitors stated income < 10.000€ in 2010 and less visitors stated income > 40.000€ in the same year (Table 2). The finding documenting the main economic crisis effect (reduction of average wages) is consistent with actual economic situation, is expected and therefore, is valid proof of respondents' credibility.

As long as income variable is concerned, the most important conclusion deriving from a second level thought is the fact that respondents give credible statements to one of the most "difficult" survey questions. Revealing income data is usually considered a taboo, an issue raising awkward and rather unpleasant feelings. It is not rare for individuals wishing to not reveal such personal information. Yet, respondents do provide relative data that, in addition, are proved realistic.

In a general socioeconomic frame of prosperity the protection of vernacular architecture holds such values that seem to evoke the positive WTP of respondents in spite of income constraints. This means that respondents of all income categories would be (or not) willing to donate money for the provision of the good. Yet, within the economic crisis, the income appears to play decisive role to stated WTP value., i.e. the lower the income, the lower the

WTP ($\chi^2= 11.384$, $df=4$, $p\text{-value}=0.023$) (Table 2). This finding comes in accordance with economic theory and shows that, even in a hypothetical setting, respondents react by taking under consideration realistic constraints. Furthermore, as the survey results show (Table 2), in 2008, the visitors of Metsovo stated a mean WTP value of 47.30€, and, in 2010, stated a mean WPT value of 17.90€. Though they are still strongly in favor of the good's provision (97.3% and 90.10% correspondingly, see Table 2) they are willing to donate much less money for it.

Table 2. Comparative data of annual income

	Annual income				
	Lower than 10.000€	10.000€– 20.000€	20.000€ – 30.000€	30.000€ - 40.000€	Higher than 40.000€
2008	9%	24%	26%	17%	24%
2010	14%	23,5%	29%	17%	17%
	SURVEYS				
	1st	2nd	3rd	4th	5th
Mean annual income	28.820€	23.580€	24.720€	24.650€	27.340€
	<i>“Yes, I believe that it should be protected and preserved”</i>				
<i>“According to your opinion, should vernacular architecture be protected and preserved?”</i>	97,30%	94%	98%	90,10%	99,20%
<i>Positive WTP</i>	41%	46,5%	72,6%	38,5%	43,8%
Mean WTP (Kaplan-Meier without zero values)	47.30€	287.90€	196.85€	17.90€	54.10€

In addition to the previous paragraph's argumentation, it should be noted that in the 2010 survey the PC elicitation format was used, which according to relative literature (Pan et al., 2012; Hoyos et al., 2010; Venkatachalam, 2004) tends to evoke higher WTP values than the OE format. Stated 2010 survey WTP values were lower, in spite of the less conservative technique used. This comes as an additional argument to the previous observation. In actual markets individuals' WTP values are constrained by income. In this hypothetical market income constraints affect WTP values.

Finally, WTP value did not exceed, in any survey, the 2% of the corresponding annual income. According to relevant literature, a 2% WTP value level is considered reasonable (Carlsson et al., 2010; Nostratnejad et al., 2016). This means, that if it were for an actual market respondents could afford such a donation, in terms of their income constraints. The fact that stated WTP values remain under 2% of income does not necessarily mean it is equal to the actual one, but it gives proof of respondents taking income constraints under consideration. Thus, it does not reflect a random answer or an unrealistic one. Indeed, it reflects a certain level of rational thinking behind it which, in turn, holds credibility.

4.2.2 The *occupational status* variable

The results of the surveys report differences among respondents' WTP value statements in relation to their occupational status. For example, individuals who stated being *employed* also stated higher positive WTP in comparison to those who stated being *unemployed*. A 54% of employed individuals stated positive WTP, while only 25% of unemployed individuals stated positive WTP. The finding suggests that individuals' current employment condition, and therefore the existence of a stable income, affected their WTP statements. Again it is proved that WTP statement does not come up as a random choice but as the result of certain considerations taken that are closely affected by income constraints. The results come in accordance with economic theory and show evidence of respondents' reliability. Given the hypothetical nature of the market, unemployed respondents could state positive WTP at higher percentages; even higher than employed ones. This, on the other hand, would not appear in an actual market. Hence, results prove evidence of actual market behavior of respondents.

Another finding was further analyzed to this direction. In the 4th survey double percentage of the respondents stated zero WTP, in comparison to the corresponding 1st survey (in 2008). The relative percentages were 20.40% and 11% correspondingly and their difference found to be statistically significant ($z=3.36 > 1.96$, $\alpha=0.05$). The finding suggests that certain limitations may have taken place leading to this percentage reduction. Since other elements remain unchanged income reduction and economic crisis hold responsibility. In accordance to economic theory (an income reduction leads to purchase restrictions) the finding shows evidence of actual market behavior.

An overall assumption, so far, is that respondents show high level of credibility in statements concerning financial elements. They also show evidence of realistic choices driven by specific economic constraints. Their attitude comes in accordance with the economic theory and, in addition, appears to simulate actual market behavior. The overall analysis shows rather no evidence of unrealistic behavior; such that could be expected in a hypothetical market.

4.2.3 The *importance of architectural heritage's protection variable*

Apart from variables concerning economic issues another set of variables was further analyzed; the one referring to respondents' attitude and beliefs on the importance of preserving local architectural heritage. Questions regarding the importance and necessity of vernacular architecture's preservation showed extremely high percentages of positive answers. Almost 90% - 98% of the respondents, in all of the five surveys, were in favor of the preservation of local vernacular architecture (Table 2). Such high and overall percentages one could expect, consequently, to have led to similarly high percentages of positive WTP; especially since payment was "hypothetical" with the meaning "not real". However, the findings of the five surveys proved this assumption wrong. The former high percentages of 90% - 98% drop to the much lower ones of 38% - 73%, correspondingly, when it comes for the WTP value statement (Table2). This practically shows that, while respondents strongly support the preservation of vernacular architecture certain constraints and limitations occur that prevent them from further stating a positive WTP in similarly high percentages. Regardless of the reasons stated in justification of the zero WTP, the finding itself suggests realism, critical thinking and honesty in respondents' statements.

One of the most appealing explanations on why respondents tend to overestimate their hypothetical WTP is their urge to increase the probability of the good's provision (Murphy et al., 2004). Thus, such strong percentages of respondents in favor of heritage's preservation would evoke high percentages of positive and rather overestimated WTP values. The finding of a very moderate positive WTP, in comparison to respondents' strong initial wish, gives additional credit to respondents' honesty.

A set of several findings of the five surveys were analyzed to the direction of testing respondents' credibility. Comparisons between four sets of questions, among the different surveys, showed that respondents' answers hold inner coherence and consistency, thus reliability. For instance:

- In Surveys 2 and 3, a certain percentage of residents claimed preservation of local vernacular architecture important because of "*its ability to attract visitors*"(14% and 6.50% in Metsovo and Sirako, correspondingly). In a different question of the surveys similar percentages of respondents appeared to have stated positive WTP justifying their answer with the same reason "*Preservation of the good will add value to the settlement as an attractive tourist destination*" (15% and 10%, in Metsovo and Sirako, correspondingly). There seems to be a consistency among different answers provided by the appearance of similar percentages in related questions.
- In Surveys 1, 4 and 5, a percentage of visitors stated positive WTP further justifying it with the statement "*vernacular architecture represents an important cultural heritage that should be preserved*". A similar percentage, in a different question

of the same survey, stated being in favor of vernacular architecture's preservation because of its *cultural and bequest value* (Table 3).

- In the 4th survey, a lower percentage of visitors would recommend Metsovo as an attractive tourist destination, in comparison with those in the 1st survey with the difference to be statistically significant ($z=3.82>1.96$, $\alpha=0.05$). On further justifying their answer, in 2010, visitors pointed out, among other reasons, the "*evident decline of local traditional character*". This justification did not exist in the 2008 survey. Also, in the 2010 survey, the percentage of visitors stating the existence of "*severe loss and decline of local vernacular architecture*" was higher than the corresponding one in the 2008 survey, with the difference to be statistically significant ($z=2.00>1.96$, $\alpha=0.05$). Repetition of similar percentages in different questions of the surveys, holding, in addition, a rational interrelation suggests evidence of realistic answers, which would be absent if responses were given randomly or without thinking.

An overall conclusion deriving from the above analysis is that the fragmental comparison of several variables can serve as testing vehicle able to reveal existence of credibility in respondents' statements. If respondents are proved of giving reliable statements to a variety of issues and present a holistic firm and consistence attitude towards the survey they cannot be then attributed with inconvenience to *some* parts of the survey. In other words, credible answers to the body of the survey are indicators of credible answers to WTP value statements as well. The opposite, lacks reliable argumentation.

Table 3. Comparative percentages between the responds to two different questions regarding the preservation of local vernacular architecture

SURVEYS	1st	4th	5th
<i>"Vernacular architecture should be preserved because it represents cultural heritage"</i>	60%	47%	60%
<i>Positive WTP justified by the phrase "it is important to preserve the cultural value of vernacular architecture"</i>	55%	46%	54%

4.3 Consistency of the CV studies with economic theory

According to relative literature (Carson et al., 2001), one indicator of CV studies' reliability is the consistency of certain results with economic theory. Among other, one such test refers to

the amount of money respondents are willing to pay for a good in comparison with its offered quantity or quality. More specifically, as the quantity or the quality of the good changes the WTP estimate also changes; for a larger amount of a good or for a better quality of it respondents are expected to pay more.

In our surveys, two different quality levels of architectural heritage were examined and valued. Respondents were asked to state WTP value for the protection of local vernacular architecture in the case of Metsovo and in the case of Sirako. Thus, two very different, in terms of built environment's quality, were under evaluation. The results of the five surveys showed that respondents state different WTP value in relation to the quality type of the cultural good. More specifically, it was proved that the higher the quality of the cultural good the higher the WTP value. Visitors of Sirako (5th survey) stated a mean WTP value which was three times higher than the corresponding one stated by the visitors of Metsovo (4th survey) (54,10€ and 17,90€, correspondingly). In addition, mean WTP value expressed as annual income's percentage in the 5th survey was two and half times higher than the corresponding one in the 4th survey (0,19% and 0,07%, correspondingly). It is worth noticing that the 5th survey took place later than the 4th survey, thus it took place in a period deeper into the general economic crisis in the country. What is of additional importance, is that the mean WTP value, in the 5th survey, during the crisis (2010-2011), was even higher than the corresponding one in the 4th survey that took place before crisis (2008) (54,10€ and 47,30€, correspondingly and 0,19% and 0,16%, correspondingly as individual's income percentage). What is suggested here, is that the cultural value of Sirako's vernacular architecture evokes significantly high economic values even in a socio-economic frame of decline. Individuals were found willing to pay higher amounts of money for a higher quality of a given good.

4.4 Comparisons between visitors' and residents' statements

Comparison among statements between residents and visitors, in each settlement, reveals some interesting results, as follows:

- Residents of Metsovo who think that local architecture is highly degraded state higher WTP (the higher the degradation the higher the WTP) ($\chi^2= 10.596$, $df=3$, $p\text{-value}=0.014$). The opposite is the case for the visitors (the higher the degradation the lower their WTP) ($\chi^2= 10.819$, $df=4$, $p\text{-value}=0.029$). The finding reveals the different relation and connection between the good and the users. Residents care more than the visitors for local architecture and this is expressed in their higher willingness to its protection. Visitors are more indifferent. If the good doesn't satisfying them they aren't willing to contribute to its protection. In addition, residents state a mean WTP

amount 16 times higher than that of the visitors in 2010 and 6 times higher in 2008 (see Table 2). This reveals the higher importance of the good for the residents.

- Residents of Sirako state a mean WTP amount that is 3.6 times higher than that of the visitors (see Table 2). In addition, while certain variables (distance of departure, decline of public space, wish of the good's preservation) ($\chi^2= 6.309$, $df=2$, $p\text{-value}=0.043$ / $\chi^2=11.047$, $df=2$, $p\text{-value}= 0.004$ / $\chi^2=6.718$, $df=1$, $p\text{-value}=0.010$, correspondingly) affect the visitors' WTP value, residents' WTP value is proved unaffected by almost any variables. The only variable affecting residents' WTP is their accordance with the law for the protection of local architecture (the more someone agrees with the law the higher the WTP value) ($\chi^2=6.703$, $df=2$, $p\text{-value}=0.035$). Findings show higher level of connection between the residents and the good, in comparison to the visitors.
- There is a steady 15% of visitors, in both settlements, who state zero WTP justifying it: *"I don't live here so I don't have to pay. Residents should pay for the protection of the good"*. The finding comes in further support of the fact that visitors appear, in general, less caring about the good, arguing that it's not their responsibility to protect it. Also, visitors often consider that residents are benefited from any deriving economic profit from the preservation of a cultural good; therefore, they should bear the cost for its protection.
- In Metsovo, WTP was found to be affected from the duration of stay, in the case of visitors, in both surveys (2008: $\chi^2=6.849$, $df=2$ $p\text{-value}=0.033$ and 2010: $\chi^2=19.399$, $df=3$, $p\text{-value}=0.000$). In specific, the longer the visitors stay the higher the WTP. In addition, in 2010, WTP was also affected by visitors willingness to revisit Metsovo in the future ($\chi^2=10.607$, $df=2$, $p\text{-value}=0.005$). In specific, those wishing to come again state higher WTP than those who won't come again in the future. The findings show that higher level of connection and familiarity with the good results in higher willingness for its protection.

All of the above findings are in accordance with relative literature (e.g. Carson et al., 1998; Montenegro et al., 2009; Garrod et al., 1996; Provins et al., 2005; Pagiola, 2001; Tuan and Navrud, 2008; Dutta et al., 2007; Kim et al., 2007; Navrud and Ready, 2002). In general, the users of a good appear to pay more for its protection than the non-users. Furthermore, residents appear more caring about their cultural goods and are more willing to contribute to their protection. Visitors usually contribute with relatively lower WTP amounts. Also, familiarity with a good and the level of its degradation increases the willingness for its protection as it comes out from surveys examining visitors WTP who either have or haven't visited the good in question.

5. Conclusions

The core importance of a CV study is indeed its hypothetical nature. In order for socially justified choices to be made a variety of possible alternatives should be presented, analyzed and evaluated. CV studies hold the advantage of capturing social preferences, beliefs, views etc., estimating the economic value of non-market goods and enabling ex-ante examination of different alternatives. Thus, CV represents a valuable tool for decision making. CV's ability to estimate non-use values when other methods fail makes it necessary and, so far, irreplaceable as Epstein (2003) points out. Hence, ongoing research on armoring the method is requisite for firm and credible economic estimations to be ensured. Nevertheless, hypothetical bias still remains at the core interest of the ongoing discussion on CV's validity. The reasons leading to WTP misstatement and the methodological approach on hypothetical bias elimination represent important issues in relative literature.

In this paper, a methodological approach on documenting reliability and credibility in WTP values, in CV studies, was presented. Based on a meta-analysis of five CV studies concerning the evaluation of vernacular architecture's preservation it is argued that hypothetical bias elimination is possible provided that inner coherence, consistency, realism and actual market simulated behavior is proved to characterize the overall body of a CV study. A variety of respondents' socioeconomic data along with a set of variables related to WTP value were analyzed. Furthermore, respondents' statements at different parts of the surveys with cross-checking questions and reliability on statements that could be tested with real data were examined. The advantage of five different studies on the same good in two different cases as well as the opportunity to examine the same good before and within the economic crisis that hit Greece enabled as to come with interesting conclusions. It came up that WTP value statements were given under consideration of specific limitations, restrictions and economic constraints. Respondents' statements revealed an inner consistent, stable and realistic behavior similar, in key-elements, to what would be expected in a real market and in accordance with economic theory's principles. We therefore argue that this kind of attitude documented should be taken as a reliable indicator of WTP values credibility, ensuring minor gap between hypothetical and actual WTP values.

The establishment of a firm methodology either eliminating or fully controlling hypothetical bias is still under research. We focused on shedding more light to the direction of turning the method's Achille's hill invulnerable. Argumentation presented in this paper, may open the field for more fertile future discussions.

References

- Ajzen I., Brown T.C., Carvajal F. (2004). Explaining the Discrepancy Between Intentions and Actions: The Case of Hypothetical Bias in Contingent Valuation. *Personality and Social Psychology Bulletin* 30(9), pp. 1108-1121.
- Arrow K.J., Solow R., Leamer E., Portney P., Radner R., Schuman H. (1993). Report of the NOAA panel on Contingent Valuation. *Federal Register* 58, pp. 4601-14.
- Azevedo C.D., Herriges J.A., Kling C.L. (2000). Ask a Hypothetical Question, Get a Valuable Answer? *Working paper 00-WP 260*.
- Bateman I.J., Munro A., Rhodes B., Starmer C., Sugden R. (1997). Does Part-Whole Bias Exist? An Experimental Investigation. *The Economic Journal, Royal Economic Society*, 107(441), pp. 322-32.
- Bedate A.M., Herrero L.C., Sanz J.A. (2012). Ex ante and ex post valuations of a cultural good. Are preferences or expectations changing?. *Journal of Environmental Planning and Management*, 55(1), pp. 127-140.
- Blackburn MK., Harrison G., Rutstrom E.E. (1994). Statistical Bias Functions and Informative Hypothetical Surveys. *American Journal of Agricultural Economics*, 75(5).
- Brown T.C., Ajzen I., Hrubes D. (2003). Further tests of entreaties to avoid hypothetical bias in referendum contingent valuation. *Journal of Environmental Economics and Management* 46, pp. 353–361.
- Carlsson F., Martinsson P. (1999). Do Hypothetical and Actual Willingness to Pay Differ in Choice Experiments? – Application to the Valuation of the Environment. Working papers in Economics, 13.
- Carlsson F., Kataria M., Krupnick A., Lampi E., Lofgren A., Qin P., Chung S., Sterner T. (2010). Paying for mitigation. A multiple county study. Discussion paper, RFF DP 10-33.
- Carlsson F., Kataria M., Krupnick A., Lampi E., Lofgren A., Qin P., Sterner T. (2013). The truth, the whole truth, and nothing but the truth—A multiple country test of an oath script. *Journal of Economic Behavior & Organization*, 89, pp. 105-121.
- Carson R.T. (2012). Contingent Valuation: A Practical Alternative when Prices Aren't Available. *Journal of Economic Perspectives*, 26(4), pp. 27–42.
- Carson R.T., Flores N.E., Martin K.M., Wright J.L. (1996). Contingent Valuation and Revealed Preference Methodologies: Comparing the Estimates for Quasi-Public Goods. *Land Economics*, 72(1), pp. 80-99
- Carson R.T., Flores N.E., Meade N.F. (2001). Contingent Valuation: Controversies and Evidence. *Environmental and Resource Economics* 19, pp. 173–210.
- Carson R.T., Mitchell R.C., Conaway M.B. (1998). Economic benefits to foreigners visiting Morocco accruing from the rehabilitation of the Fes Medina, World Bank study.

- Champ P.A., Bishop R.C. (2001). Donation Payment Mechanisms and Contingent Valuation: An Empirical Study of Hypothetical Bias. *Environmental and Resource Economics* 19, pp. 383–402.
- Champ P.A., Moore R., Bishop R.C. (2009). A comparison of approaches to mitigate hypothetical bias. *Agricultural and resource economics Review*, 38(2), pp. 166-180.
- Cummings R.G., Taylor L.O. (1999). Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method. *The American Economic Review* 89(3), pp.649-665.
- Damigos, D., Kaliampakos, D. (2006). The “battle of gold” under the light of green economics: A case study from Greece. *Environmental Geology*, 50(2), pp. 202 – 218.
- Dutta M., Banerjee S., Husain Z. (2007). Untapped demand for heritage: A contingent valuation study of Prinsep Ghat, Calcutta. *Tourism Management* 28(1), pp. 83–95.
- Epstein R.A. (2003). The Regrettable Necessity of Contingent Valuation. *Journal of Cultural Economics* 27, pp. 259–274.
- Garrod G.D., Willis K.G., Bjarnadottir H., Cockbain P. (1996). The non – priced benefits of renovating historic buildings. A case study of Newcastle’s Grainger Town. *Cities*, 13(6), pp. 423-430.
- Giannakopoulou S. (2012). Evaluation of vernacular architecture of mountainous areas with the application of Environmental Economics methods. PhD Thesis, National Technical University of Athens (*in Greek*).
- Giannakopoulou S., Damigos D., Kaliampakos D. (2011). Assessing the Economic Value of Vernacular Architecture of Mountain Regions Using Contingent Valuation. *Journal of Mountain Studies*, 8, pp. 1-9.
- Giannakopoulou S., Kaliampakos D. (2012). Cultural Economics: Reliable for protecting architectural heritage. 17th ACEI International Conference, Kyoto, Japan.
- Giannakopoulou S., Kaliampakos D. (2012). Restoration of traditional architectural character. A cost-benefit analysis in a Greek mountainous town, based on the implementation of CVM. HERITAGE Conference, Porto, Portugal.
- Hanemann W.M. (1994). Valuing the Environment through Contingent Valuation. *Journal of Economic Perspectives*, 8(4), pp.19-43.
- Harrison, G.W., Rutström E.E. (2008). Experimental Evidence on the Existence of Hypothetical Bias in Value Elicitation Methods. In: Handbook of Results in Experimental Economics, ed. C. Plott and V.L. Smith, New York: Elsevier Science.
- Hoyos D., Mariel P. (2010). Contingent Valuation. Past, Present and Future. Prague Economic Papers 4.

- 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. *Tourism Management* 28(1), pp. 317–322.
- Kling C., Phaneuf D., Zhao J. (2012). From Exxon to BP: Has Some Number Become Better than No Number? *Journal of Economic Perspectives*, 26(4), pp. 3-26.
- List J.A. (2001). Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards. *American Economic Review* 91, pp. 1498-1507.
- List J.A., Gallet C. (2001). What Experimental Protocol Influence Disparities between Actual and Hypothetical Stated Values? *Environmental and Resource Economics*, 20, pp. 241-254.
- Little J. and Berrens R. (2003). Explaining Disparities between Actual and Hypothetical Stated Values: Further Investigation Using Meta-Analysis. *Economics Bulletin* 3 (6), pp. 1–13.
- Loomis J. (2011). What's to know about Hypothetical Bias in Stated Preference Valuation studies? *Journal of Economic Surveys*, 25 (2), pp. 363–370.
- Mitchell R.C., Carson R.T. (1989). Using Surveys to Value Public Goods: The Contingent Valuation Method. John Hopkins Press, Baltimore.
- Mohammed E.Y. (2012). Contingent valuation responses and hypothetical bias: mitigation effects of certainty question, cheap talk, and pledging. *Environmental Economics*, 3(3).
- Montenegro A.B., Huaquin M.N., Prieto L.C.H. (2009). The valuation of historical sites: A case study of Valdivia, Chile. *Journal of Environmental Planning and Management* 52(1), pp.97-109.
- 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. *Environmental and Resource Economics* 44, pp. 307–326.
- Murphy J.J., Allen P.G., Stevens T.H., Weatherhead D. (2005). A Meta-Analysis of Hypothetical Bias in Stated Preference Valuation. *Environmental and Resource Economics*, 30, pp. 313–325.
- Murphy J.J., Stevens T.H. (2004). Contingent Valuation, Hypothetical Bias and Experimental Economics. *Agricultural and Resource Economics Review* 33(2), pp. 182-192.
- Navrud S. and Ready R.C. (eds) (2002). *Valuing cultural heritage: Applying environmental valuation techniques to historic buildings, monuments and artefacts*. Edward Elgar Publishing Ltd., UK.
- Noonan D. (2003). Contingent valuation and cultural resources: a meta-analytic review of the literature. *Journal of Cultural Economics*, 27, pp. 159-176.

- Nostratnejad S., Rashidian A., Dror D.M. (2016). Systematic review of willingness to pay for health insurance in low and middle income countries. *PloS one* 11(6), e0157470.
- Pagiola S. (2001). Valuing the Benefits of Investments in Cultural Heritage: The Historic Core of Split. Presentation at the International Conference on Economic Valuation of Cultural Heritage, 19-20 October, Cagliari.
- Pan J., Zeng W. and Zhang W. (2012). The Willingness to Pay for the Ecological Compensation of Min River Basin -- Based on the Survey of Chengdu 282. *Households Environment and Pollution*, 1(2).
- Provins A., Pearce D., Ozdemiroglu E., Mourato S., Morse-Jones S. (2005). Valuation of the historic environment: The scope for using economic valuation evidence in the appraisal of heritage – related projects. *Progress in Planning*, 69, pp.131 – 175.
- Rizzo I., Throsby D. (2006). Cultural heritage: economic analysis and public policy. In: Victor A. Ginsburgh and David Throsby (eds.), *Handbook of the Economics of Art and Culture*, pp. 984–1016. Amsterdam: Elsevier/North-Holland.
- Samnaliev M., Stevens T.H., More T. (2006). A comparison of alternative certainty calibration techniques in contingent valuation. *Ecological Economics* 57, pp. 507-519.
- Santagata W., Singorello G. (2000). Contingent Valuation of a Cultural Public Good and Policy Design: The Case of “Napoli Musei Aperti”. *Journal of Cultural Economics* 24, pp. 181–204.
- Tentes, G., Damigos, D. (2012). The Lost Value of Groundwater: The Case of Asopos River Basin in Central Greece. *Water Resources Management* 26, pp. 147–164.
- Throsby D. (2000). *Economics and Culture*. Cambridge University Press, Cambridge, UK.
- Tuan T.H., Navrud S. (2008). Capturing the Benefits of Preserving Cultural Heritage. *Journal of Cultural Heritage* 9(3), pp. 326-337.
- Veisten K., Navrud S. (2006). Contingent valuation and actual payment for voluntarily provided passive-use values: Assessing the effect of an induced truth-telling mechanism and elicitation formats. *Applied Economics* 38, pp. 735–756.
- Venkatachalam L. (2004). The contingent valuation method: a review. *Environmental Impact Assessment Review* 24, pp. 89–124.
- Vossler C.A., Kerkvliet J. (2003). A criterion validity test of the contingent valuation method: comparing hypothetical and actual voting behavior for a public referendum. *Journal of Environmental Economics and Management* 45, pp. 631–649.
- Xie J., Gao Z. (2013). The Comparison of three Non-hypothetical Valuation Methods: Choice Experiments, Contingent Valuation and Experimental Auction. Selected Poster prepared for presentation at the Southern Agricultural Economics Association (SAEA) Annual Meeting, Orlando, Florida, 3-5 February.