This paper has not been submitted elsewhere in

identical or similar form, nor will it be during the first three months after its submission to the Publisher

Controlling for biases in primary valuation studies: a meta-analysis of international coral reef values

Abstract

This paper updates an international database on coral reef valuation studies, conducts a meta-analysis with a focus on treatments in primary studies that are intended to mitigate biases in the stated preference valuation methods and evaluates the ability of the meta-analysis function to make international benefit transfer predictions. This paper evaluates the common biases identified by the National Oceanic and Atmospheric Administration's (NOAA) guidelines, as well as other factors in the context of different income levels of economies. The results show that common bias treatments are associated with variations in willingness-to-pay estimates among distinct economies: high-income economies are likely to "cushion" the bias effects better than upper-middle, lower-middle and low-income economies. Indeed, the influence of these biases, particularly among distinct economies, suggests that international benefit transfers require more attention when transferring values from high to low economies, particularly when many low-income economies are dependent upon coral reefs.

Key words

Biases, coral ecosystem, meta-analysis, valuation, willingness to pay

JEL

Q54, Q57

Abbreviations

Benefit transfer (BT)

Breusch–Pagan Lagrange multiplier (BPLM)

Choice experiment (CE)

Contingent valuation method (CVM)

Convention on biodiversity (CBD)

Gross domestic product (GDP)

International Monetary Fund (IMF)

Marine protected area (MPA)

Meta-regression analysis (MRA)

National Oceanic and Atmospheric Administration (NOAA)

Ordinary least square (OLS)

Purchasing power parity (PPP)

Random-effects model (REM)

Travel cost method (TCM)

Willingness to pay (WTP)

World Development Indicator (WDI)

1. Introduction

Living coral ecosystems provide a wide range of goods and services to marine biodiversity as well as to society. However, such delicate biomes are on the verge of being destroyed, both by natural and by anthropogenic events. Several of these essential services include regulating, provisioning, cultural and supporting services. The case of coral reef management is important, as pointed out by the Intergovernmental Panel on Climate Change (1995), in that these resources are expected to be among the first environments to be stressed by climate change. Moreover, apart from the threat of climate change, other direct threats stem from anthropogenic activities (e.g., overuse and pollution). Consequently, current economic research interests in coral reefs are not only the measurement of non-use and use values, but also the evaluation of alternative management strategies for coral reef sites.

In recent years, the focus of non-market valuation studies has been on the use of utility-theoretic models of valuation, including the use of hypothetical behaviours, as in stated preference methods of contingent valuation and choice experiments, and observed behaviours, as in revealed preference methods such as hedonic property and travel cost models. The majority of primary studies evaluate cultural service flows from coral reefs, with few primary studies evaluating regulating, provisioning and supporting services. Specifically, cultural services have been measured in the context of non-extractive recreation uses in the form of tourism.

Given that non-market values are important sources of information when assessing coral reef policies, management strategies and marine spatial planning, the use of methods that provide credible, reliable and valid estimates is paramount in the development of non-market valuation methods. Many sources of bias have been identified in the non-market valuation literature, in particular for stated preference methods (Mitchell and Carson 1989; Bateman et al. 2002). Some biases may be inherent in an individual's perception of social interactions; however, in many cases, the researcher can design a study that collects primary data in a way that minimizes the potential biases. This is in part the motivation behind the National Oceanic and Atmospheric Administration's (NOAA) Blue Ribbon Panel recommendations for conducting contingent valuation studies (Arrow et al. 1993). When biases are mitigated, the confidence in the validity and reliability of the survey instruments and the resultant value estimates increases. Hence, reducing common biases has important policy implications not only for accurately estimating values at primary study sites, but also when using the outcomes from these studies to inform related management and policy decisions at other sites (i.e., through benefit transfer). For example, the NOAA's natural resource damage assessment rules suggest reducing stated preference value estimates by a factor of two to account for hypothetical bias (List and Gallet 2001).

Hypothetical bias is one of the more broadly recognized issues in stated preference techniques in non-market valuation when the respondent overstates his/her willingness-to-pay (WTP) value. One reason attributed to this bias is the familiarity of the good to the true state of WTP values (Cherry and Shogren 2007).This form of bias manifests as higher expressions of value when derived from a stated preference survey than is found in actual cash or comparable revealed preference elicitations, thus leading to the NOAA calibration rule as well as instrument and statistical calibrations. Other research has suggested alternative approaches to dealing with hypothetical bias in stated preference surveys. Another *ex post* method is to use a certainty equivalent follow-up question to the valuation question. Two primary certainty follow-up methods have been used—a numerical scale that typically consists of a rating of certainty of payment on a 10-point Likert scale (Champ et al. 1997) and a polychotomous choice consisting of selecting from a set of responses (e.g., definitely yes, probably yes, maybe yes, maybe no, probably no, definitely no) (Ready et al. 1995). Neither of these follow-up certainty methods has been found to be superior, and while they both help mitigate hypothetical bias, they also result in less precise estimates (Akter et al. 2008).Furthermore, certainty follow-ups provide *ex post* calibration factors or thresholds of inclusion, but do not necessarily mitigate hypothetical bias in the survey design.

Researchers also have investigated *ex ante* methods to reduce hypothetical bias. One of these methods is the use of a cheap talk script (i.e., a script inserted into a survey prior to the valuation questions that explicitly acknowledges, for the respondents, the reality of hypothetical bias) with the intention that if the respondents are made aware of this form of bias, they can be instructed to minimize it or avoid it altogether. Since Cummings and Taylor's (1999) introduction to and empirical success with a cheap talk script, other researchers have tested the method and found it to work in some, but not all, cases (Murphy et al. 2005b; Aadlund and Caplan 2006). Another *ex ante* approach to mitigating hypothetical bias is suggested through the use of consequential preference questions, in particular choice-based questions (Carson and Groves 2007). Inconsequential questions are not perceived by the respondent to have any actual effects on policy or management outcomes, thus there is little incentive for the respondent to provide a precise and consistent estimate of his/her value. In this regard, questionnaire design is important to avoid situations known to increase the hypothetical nature of responses.

The objectives of this paper are 1) to update and expand on a previous global database of coral reef valuation studies; 2) to test the NOAA (Arrow et al. 1993) guidelines and hypothetical bias mitigating strategies instated preference studies; and 3) to evaluate the transferability of value predictions from the meta-analysis regression function to other sites, particularly in developing countries where coral reef resources are abundant and facing natural and anthropogenic threats. This is the first meta-analysis to focus on the NOAA Blue Ribbon Panel's guidelines specifically related to procedures and bias mitigating strategies for primary stated preference valuation studies conducted outside of laboratory experiments.

The remainder of this paper is structured as follows: Section 2summarizes the previous meta-analyses of coral reef valuation studies and hypothetical bias studies, comparing them with the current meta-analysis. Section 3 introduces the database and a data description of the variables of interest and Section 4 reports the empirical results. A discussion of the results and the conclusion are presented in Section 5.

2. Previous meta-analyses of coral reef values and hypothetical bias

We evaluate the effects of bias mitigation strategies in the literature at the aggregate level through the use of meta-regression methods. Meta-regression analysis (MRA) is the statistical analysis of a body of research (Stanley 2001; Nelson and Kennedy 2009): in this case, coral reef valuation studies. MRA has been applied to valuation research outcomes in areas such as groundwater quality (Boyle et al. 1994), cultural resources (Noonan 2003), wetland resources (Woodward and Wui 2001), outdoor recreation (Rosenberger and Loomis 2000) and endangered species (Loomis and White 1996). Nelson and Kennedy (2009) evaluated 140 MRAs in the fields of environmental and resource economics, culminating in recommended best practices for these types of analyses. While MRA is primarily conducted to synthesize a body of research and test hypotheses, it also has the potential to predict values for unstudied sites and contexts (i.e., benefit transfer).

The estimated meta-regression function may be adjusted based on relevant measures of the policy site, predicting valid and reliable estimates of value, as well as testing and mitigating selection effects within the literature (Johnston and Rosenberger 2010). However, MRA is not a panacea for benefit transfers, being itself limited by the inherent errors, omissions and transferability of primary research outcomes (Rosenberger and Stanley 2006; Rosenberger and Johnston 2009). The realm of international benefit transfers, whether via MRA, point estimates or other function transfers, remains an open area of enquiry with evaluations showing both successes and failures in the process (Johnston and Rosenberger 2010).

Two previous MRAs have focused on coral reef valuation studies across the globe, each providing insights and foundations on which to build the coral reef MRA of this paper. Brander et al.'s (2007) MRA evaluated33 studies, providing a total of 73 observations on coral reef values derived in primary research using a variety of utility-theoretic and engineering valuation approaches, with contingent valuation and travel cost methods dominating the total number of value estimates in the database. They estimated the log-level form of a random-effects, multi-level modelling approach that accounts for intra-level correlations defined by primary study authors, finding a significant author effect in this literature. The data were further expanded by measuring GDP and population density within 50km of each study site, although these added exogenous regressors were not significant in the final models. The final model included factors associated with location (i.e., dive site area, East Africa), visitation (i.e., number of visitors to the study site), activities (i.e., snorkelling) and valuation methods (i.e., travel cost, production function, net factor income and gross revenue, with contingent valuation being the omitted category). Other categories within these attributes, along with other attributes, were not found to be statistically significant and therefore were not included in the final model.

All the valuation methods resulted in statistically higher-value estimates than contingent valuation studies, with the exception of gross revenue, which showed a positive shift from contingent valuation, but was insignificant in the final model. Brander et al. (2007) also applied an n-1 jackknife out-of-sample evaluation of benefit transfer accuracy. Based on 73 out-of-sample tests, the mean transfer error was 186% and the median transfer error was 79%. Based on these transfer error tests, they concluded that MRA benefit transfers, while accurate in some contexts, are limited in part by the lack of high-quality primary studies in other contexts.

Londoño and Johnston (2012) evaluated the coral reef valuation literature with primary foci on ensuring data uniformity (i.e., reducing heterogeneity through metadata sample selection and welfare consistency), comparing classical with Bayesian MRA and augmenting primary study data with external coral reef biophysical data. Thus, their MRA evaluated 29 studies, providing a total of 71 observations on coral reef values derived in primary research using the contingent valuation method.[1] They also estimated the log-level form of a random-effects, multi-level modelling approach, but using each study as a potential source of intra-level correlations, intra-study variances were not significant in their models. The data were expanded using an external coral reef database from which measures of habitat size, marine protected area status and reef quality were derived for each included study site. The final model included factors associated with primary study methods (i.e., value elicitation type, on-site survey mode), location attributes (i.e., East Africa), activities (i.e., snorkelling and diving) and resource characteristics (i.e., reef area, marine protected area, reef quality and reef type). They found that methods and resource characteristics were significant moderators of variations in coral reef value estimates within the contingent valuation literature. Moderator variable partial effects were consistent between the classical and the Bayesian models, although the latter models provided more precise estimates.

Londoño and Johnston (2012) also applied an n-1 jackknife out-of-sample evaluation of benefit transfer accuracy. Based on 71 out-of-sample tests, the mean transfer error was 94% and the median transfer error was 43%—effectively half of the transfer error found by Brander et al. (2007). They concluded that the addition of resource characteristics measures (e.g., marine protected area status and live coral coverage) and limiting the analysis to utility-theoretic valuation methods improved the ability of meta-regression function-based international benefit transfers, although additional research is needed on coral reefs in general and international benefit transfers in particular.

Meta-regression analysis has been used to evaluate hypothetical bias in the non-market valuation literature. List and Gallet (2001) evaluated 29 studies, providing 58 estimates of calibration factors (i.e., mean hypothetical value divided by mean actual value); each study derived both hypothetical and actual value estimates in an experimental setting. The context of these experiments ranged from students to the general public, lab to field experiments and private to public goods, and used a variety of elicitation techniques including auction methods, open-ended and dichotomous choice. Based on various model assumptions, their MRAs explained 40–50% of the variation in calibration factors. Significant moderator effects were estimated for willingness to pay (lower calibration factors relative to willingness to accept) and private goods (lower calibration factors relative to public goods), along with reductions in calibration factors associated with two auction techniques. The dichotomous choice-based elicitation, while not statistically significant, was negative. In general, List and Gallet (2001) suggested that the mean calibration factor is about three, not two, as suggested by the NOAA's natural resource damage assessment rule.

Murphy et al. (2005a) updated the List and Gallet (2001) meta-analysis by focusing on willingness-to-pay estimates only (i.e., a few studies estimated willingness to accept, which were substantially larger than their willingness-to-pay counterparts) and including only studies that used the same elicitation method for both hypothetical and actual value estimates (i.e., to control for differences in elicitation methods within studies). Their meta-analysis consisted of 28 studies, providing 77 estimates in the MRA. Consistent with List and Gallet (2001), Murphy et al.'s (2005a) mean calibration factor for their metadata was 2.6; however, due to the positively skewed distribution of calibration factors, they focused on the median estimate of 1.35.

The metadata were modelled with the dependent variable as the natural log of the actual value, and the independent variables including a quadratic natural log of hypothetical value and other moderator variables. The MRA explained about 90% of the variation in the natural log of the actual value. In general, they found that group settings, like lab experiments, resulted in relatively higher levels of hypothetical bias than individual responses to surveys, and that choice-based elicitation methods (e.g., dichotomous and polychotomous choice, referendum, payment card and choice experiments) resulted in lower levels of hypothetical bias than other elicitation methods (e.g., open-ended and auction mechanisms). Private goods were also found to have lower levels of hypothetical bias associated with them.

This paper builds on these prior works by: 1)adding more recent coral reef valuation studies; 2) expanding the database to include both natural and artificial reefs; and 3) evaluating intra-study biases associated with applied valuation survey research conducted on a natural resource. Furthermore, given that the contingent valuation method (CVM) has a greater number of potential biases associated with it than other valuation methods, such as choice experiments, hedonic property methods and travel cost methods, we also estimate a CVM-only MRA to isolate the effects within applications of this method to coral reef valuation.

3. Data collection and variable selection

The identification of empirical coral reef valuation studies began with Brander et al.'s (2007) bibliography. This list of studies was then supplemented with an extensive search of several databases, such as Econlit, International Bibliography of Social Sciences and Web of Knowledge, Environmental Valuation Reference Inventory (EVRI), Environmental Valuation, Cost-Benefit News (EVCBN), ReefBase, Nature Valuation and Financing Network Case Study Database and Google Scholar. An effort was made to include grey literature, such as conference proceedings, unpublished academic papers and vendor reports, to ensure broad coverage of empirical studies and minimize potential biases associated with metadata sample selection (Rosenberger and Johnston2009).

The studies included in the database met the following criteria: 1) they provided original estimates of coral reef values in monetary units on a per person basis (i.e., willingness to pay); 2) they provided estimates for either natural or artificial reefs; 3) the reporting in the study was in the English language; and 4) they provided sufficient information on the key moderator variables. The application of these rules resulted in 53 studies, providing 164 estimates of coral reef values, making it twice as large as Brander et al.'s (2007) study. The year data were gathered across primary studies ranging from 1993 to 2010. The appendix table provides a summary of studies by author, year, methodology and region.[2] Most studies estimated recreation values elicited from local and non-local tourists.CVM is the most commonly used stated preference approach, with around 70% of the total studies, followed by choice experiments (CE; 20%) and travel cost methods (TCM; 12%). Table 1 shows the number of studies conducted between the years 1993–1998, 1999–2004 and 2005–2010 by different regions, in which the number of CEs has increased, especially among small islands and emerging economies, whereas CVM is being pursued actively regardless of region.

Table 1: Selected number of stated preference observations by group (year) and region

Year/Region	**Stated preference type**	
1993–1998	CE	CVM
Caribbean and small islands	0	1
N. America, US	0	1
1999–2004		
Africa, East	0	2
Caribbean and small islands	4	11
Europe, South	5	1
N. America, US	0	44
Oceania, Mela, Micro and Poly	0	6
South-East Asia	0	11
2005–2010		

Africa, East	0	2
Australia	6	0
Caribbean and small islands	3	4
Central America	0	2
Europe, South	8	3
N. America, US	1	3
Oceania, Mela, Micro and Poly	3	7
South Asia	0	2
South-East Asia	0	12

Table 2 summarizes the variables used in the analysis, including the dependent variable, which is the annual mean willingness to pay per person (US$ 2005) adjusted for inflation rates using the implied purchasing power parity (PPP) rate found in the International Monetary Fund (IMF) World Economic Outlook 2006. Moderator or independent variables were either coded from each study's documentation following prior meta-regression analyses, NOAA guidelines and hypothetical bias literature or derived from external, secondary sources.

Table 2: Variables of interest according to the NOAA, value elicitation and context factors

Variables	Details	Obs.	Mean	Std dev.	Min.	Max.	Exp. sign
Variables associated with NOAA general guidelines							
Sample size	1=sample size greater than 100; 0=otherwise	131	0.8550	0.3535	0	1	+
Interview survey mode: personal	1=in-person or phone interview; 0=otherwise	131	0.6870	0.4655	0	1	+/-
Variables associated with NOAA value elicitation guidelines							
WTP	log form of WTP per person/year	131	1.2740	3.2840	-7.4610	13.230	-
Referendum format	1=close-ended choice question; 0=otherwise	131	0.3740	0.4857	0	1	+/-
Picture or photo	1=used photographs/pictorials in survey; 0=otherwise	131	0.3130	0.4655	0	1	+/-
Substitutes: reminder of undamaged substitutes	substitutes if mentioned	131	0.4351	0.4977	0	1	-

Cheapscript	1=cheap talk script used; 0=otherwise	131	0.5038	0.5019	0	1	-
Certainty	1=certainty follow-up used; 0=otherwise	131	0.1985	0.4004	0	1	-
Additional factors							
TCM	1=revealed preference, travel cost method; 0=otherwise	131	0.0992	0.3001	0	1	+/-
CVM	1=stated preference, contingent valuation method; 0=otherwise	131	0.6794	0.4685	0	1	+/-
CE	1=stated preference, choice experiment method; 0=otherwise	131	0.2061	0.4061	0	1	+/-
Mail	1=mail survey; 0=otherwise	131	0.1603	0.3683	0	1	+?
On-site	1=on-site survey; 0=otherwise	131	0.4351	0.4977	0	1	-?
Open-ended	1=stated preference, open-ended value elicitation mode; 0=otherwise	131	0.1069	0.3101	0	1	+/-
Payment card	1=stated preference, payment card value elicitation mode; 0=otherwise	131	0.1985	0.4004	0	1	+/-
Entrance fee	1=entrance fee is payment mechanism; 0=otherwise	131	0.4275	0.4966	0	1	-
Translation	1=questionnaire translated to or from English; 0=otherwise	131	0.2290	0.4218	0	1	+?
International users	1=sample consisted of international tourists; 0=otherwise	131	0.2443	0.4313	0	1	+

Year	year of data collection, converted to index by years past 1991	131	11.229	3.9274	0	18	+/-
Publication	1=peer-reviewed publication; 0=otherwise	131	0.4198	0.4954	0	1	+/-
Natural reef	1=natural reef; 0=otherwise	131	0.8397	0.3683	0	1	+?
Marine protected area (MPA)	1=MPA; 0=otherwise	131	0.5115	0.5018	0	1	+
GDP per capita	in thousands (constant 2005 international $) reported by World Bank (2010) WDI database	131	17.290	16.723	0.3478	56.459	+
Coral/MPA	ratio of reef area to MPA (sq km)	131	4.9117	21.937	0.0037	165.18	+/-?
CBD	1=signatory to Convention on Biodiversity Diversity (CBD) in 1992–1993; 0=otherwise	131	0.6260	0.4857	0	1	+

The NOAA guidelines (Arrow et al. 1993) are categorized as either general guidelines or value elicitation guidelines. The primary context of these guidelines is natural resource damage assessments; therefore, they are not broadly relevant to all coral reef valuation studies (e.g., adequate time lapse since damages, temporal averaging of responses). For other guidelines, credible proxy measures are either not available (e.g., conservative design) or not reported (e.g., testing of interviewer effects, pretesting survey instrument and other survey components). The variables associated with some of the NOAA's general guidelines include: sample size, response rate, interview mode, referendum format and the use of pictures and substitutes. Moreover, other variables associated with hypothetical bias calibration were included, like certainty levels and cheap talk scripts. Other additional variables considered include: all valuation methods, elicitation mode, respondent type, peer-review status, etc. These data were further augmented with information derived from secondary sources, including context characteristics such as GDP per capita, coral reef area and marine protected area (MPA).

As mentioned in the NOAA guidelines, the sample size and type matter. The sample size may be a proxy measure for several effects, namely larger sample sizes may be more representative of a population of interest, *ceteris paribus*, than smaller sample sizes. The precision of effects (i.e., standard errors) is a function of sample size, thus larger sample sizes, *ceteris paribus*, may lead to more precise estimates. Additionally, small sample sizes may be indicative of small budgets. In our case, we used the sample size as the studies did not provide detailed information on the sample-specific designs. Instead, we wanted to find out whether a sample size greater than 100 had any impact on the meta-analysis. Our a priori expectation is that larger sample sizes will lead to larger WTP values given that smaller price effects can be measured (i.e., price effects are commonly in the denominator of welfare formulas).

Other general guidelines include the response rates (non-response bias—representativeness); the response rate is considered important as high non-response rates would imply that the sample results are unreliable. In our case, the number of studies that reported the response/non-response rates was too low to predict the effect of this variable. Additionally, we focused on the survey mode (interview—in-person, phone, mail and Web) as well as the elicitation format (closed-ended, open-ended, payment card, etc.).

Variations in WTP estimates are often found to be associated with different elicitation formats when evaluated through the use of meta-regression analysis. As Johnston and Rosenberger (2010) noted, primary studies have shown WTP to vary in theoretically expected and predictable ways. Furthermore, Gallet and List (2001) and Murphy et al. (2005a) found choice-based formats to be associated with lower hypothetical bias than other methods. Not only have these different formats led to differences in the estimated WTP values (Carson et al. 2001), but they may also have imposed different levels of cognitive burden on the respondents. For instance, respondents typically find open-ended questions more strenuous to answer than closed-ended questions (Hanemann 1994). Therefore, in addition to identifying closed-ended formats, as recommended by the NOAA Panel, dummy variables were included for open-ended and payment card formats. Notwithstanding the potential hypothetical bias issues associated with the elicitation format, the valuation method has been found to have measurable differences, including revealed preference (travel cost method) and stated preference (CVM, CE) (Rosenberger and Loomis 2001).Other valuation approaches were identified, such as CVM, travel cost methods (TCM), choice experiment (CE), expenditure and gross revenue.

With regard to the interview survey mode, the NOAA recommends the use of personal interviews; in other words, face to face as well as telephone compared with mail. In our analysis, we included personal as well as mail and Internet (Web) in the final analysis. According to the NOAA, personal interviews are to be encouraged as one can motivate respondents with the use of graphic illustrations. However, the guidelines suggest pretesting photographs, maps, etc., not only because they may be instrumental in conveying information about a natural resource, but also because photos may influence WTP in other ways. Mitchell and Carson (1989) noted that the use of pictures to present actual scenarios runs the risk of “reality enhancement” and respondents may provide WTP estimates based on their subjective views of the pictures, not of the policy or management context being evaluated. Given that little information is contained in the primary studies regarding their pretesting for bias effects from photos, we identified whether the primary study used photos or not in the questionnaire. It is expected that photographs may have mixed impacts on the respondents, such as positive or negative feelings.

The NOAA also identifies that respondents should be aware of substitute sites and alternative goods; failing to consider other public or private goods could result in overestimates of WTP. With regard to willingness to accept (WTA) vis-à-vis WTP, the NOAA guidelines point out that the former is less used in valuing the reduction of future damage than the latter because it produces unreasonable estimates. Furthermore, the NOAA guidelines suggest that the difference between the two approaches is trivial, though WTA was found to be important by List and Gallet (2001).Evidently, this database contains only WTP estimates as most primary studies requested respondents to pay for the coral-reef-related services, which is considered a normal phenomenon in recreation and/or tourism activities.

Hypothetical bias mitigation strategies include *ex ante* methods, such as a cheap talk script and the use of choice-based elicitation formats. *Ex post* methods include the use of certainty follow-up questions or some other form of calibrating estimated values. Additionally, the use of a cheap talk script in the questionnaire design has been suggested and found to reduce hypothetical bias (Cummings and Taylor 1999; List and Gallet 2001; Murphy et al. 2005b; List et al. 2006).Cheap talk is the use of script statements that caution respondents on biases, prior to providing their WTP values for a good (or service). A cheap talk script was identified, in particular, when respondents were reminded about their expenditure/budget constraints. Other details of scripts used, such as the length and content of these statements, may matter, but this information was not reported in sufficient detail to be coded for the MRA. Surveys that included scripts are expected to report lower WTP values given that they reduce the hypothetical bias.

The use of certainty follow-up questions has been found to reduce hypothetical bias, as previously noted (Akter et al. 2008),although only if the follow-up questions are used to truncate the data or adjust them based on a calibration factor. Thus, we code the use of a certainty follow-up question to identify whether there is a systematic effect on coral reef valuations. However, the level of certainty reported by respondents is not reported consistently; we can only evaluate whether the use of certainty follow-up questions resulted in reduced WTP estimates reported by the studies that use them.

The payment vehicle is another important component of the valuation question design, including options such as payments in the form of increased taxes, utility bills, general prices of goods purchased and fees for the goods or service (Hanley and Spash1993). In this study, entrance fees are coded as a payment vehicle apart from other forms of payments. This is because entrance fees are typical payment modes for recreation services in tourism. Entrance fee payments are also more consequential in that they take into account the effect of an on-site survey as a pertinent variable in recreation valuation as the respondents are in the presence of the goods and/or services being evaluated. Furthermore, imposing entrance fees is one plausible method for a recreation good or service to generate revenue for damage mitigation or other reinvestments in the resources, goods and services delivered. Moreover, we considered the NOAA recommendation that such payment vehicles should preferably be applied to referendum situations. In our case, the interaction term, namely entrance fees, with closed-ended questions suggested that the effect on WTP estimates is positive because respondents are generally comfortable and understand the task of closed-ended as opposed to open-ended questions, when using a payment vehicle like fees.

The translation of a questionnaire may also influence WTP value expressions, especially when describing hypothetical scenarios. However, as pointed out by Bulmer and Warwick (1993), other challenges[3]are encountered when translating the surveys into local languages; however, these issues are not covered in this study. A translation dummy variable is also included to examine whether the questionnaire being translated to or from English for the respondents may affect the WTP estimates.

In this case, we expect translation to other languages to lower the WTP estimates as the information provided to the respondents is less complicated cognitively with regard to the valuation context.

Other factors coded include whether the survey was strictly for non-resident respondents, who are presumed to be international tourists. In this case, the extension of the valuation market suggests that non-resident respondents will report higher WTP values than residents. Moreover, other variables of interest include on-site valuation, in which WTP estimates are expected to vary for those valuing the reef. On-site, for example, expressed values might be lower or higher depending on the experience the respondents had with the coral reef.

As found in other meta-analysis literature, publication in a good-quality journal—in other words, a peer-reviewed publication—may imply higher quality of the valuation work, indirectly signifying that the sample framework and/or biases have been reduced. However, the effects maybe either positive or negative depending on the sample size. In other words, screening may lead to higher sample sizes, signifying positive estimates or more conservative methods affecting WTP estimates negatively. Rosenberger and Johnston (2009) also showed that values may differ based on whether they are published in a peer-reviewed outlet or not, although a statistically significant difference is not always found across different literatures.

Additionally, we introduced the data year to check whether there is a trend in coral reef values. Rosenberger and Johnston (2009) showed that trend effects are mixed in a variety of environmental valuation data. The year of data collection, converted to an index of the number of years past 1991, implies that a value trend greater or less than inflation may also be correlated with other changes in research methods and site types over time.

Natural and artificial reefs are included and coded in the analysis. Mixed effects on WTP estimates are expected; in one case, the WTP may be larger for natural reefs due to higher passive use values, although as reported by Ghermandi et al.'s (2010) meta-analysis of wetland valuation studies, constructed wetlands had systematically higher value estimates than natural wetlands due to their location, implied need, etc.

The data were further augmented with information derived from secondary sources, including country-specific context characteristics such as gross domestic product (GDP) (i.e., a proxy for income) as reported by the World Bank (2010) World Development Indicator (WDI) database and converted using PPP (constant 2005 international $). It is expected that GDP positively influences WTP estimates as higher incomes increase the ability to pay and values.

The inclusion of the ratio of reef area to MPA (sq km), as reported in World Databases on Protected Areas (2009), implies that the WTP values may have mixed effects; in other words, positive WTP may signal high-quality, unique resources in the area, or negative WTP that the availability of substitutes is prevalent. Another inclusion is being a signatory to the Convention on Biodiversity (CBD); we expect the WTP estimates to be positively associated with the site; in other words, it should signal higher-quality or unique resources at country or regional levels.

Other relevant variables, as related to the NOAA but excluded from the final analysis due to aninability to record (i.e., not reported in primary studies) or multicollinearity with other variables, include: pretest, questionnaire design characteristics, protest size and inclusion of splitsamples, etc.

4. Empirical model and discussion of results

For the meta-regression analysis (MRA), the semi-log form is used (see equation 1 for ordinary least square (OLS)) in which the log WTP per person/year (dependent) is regressed against three independent components: methodological (or study) characteristics such as valuation type, elicitation format, year, cheap talk script, sample size, publication type, vehicle payment; site (or good) characteristics including reef to MPA ratio, on-site survey, user type; and context (or activity) associated with GDP per capita as a proxy for income and CBD year.

$$y_1 = \alpha + \beta' x_1 + {}_{+}\varepsilon_1 \qquad (1)$$

where y is the log of WTP per person/year adjusted to 2005 US$ and α and β are the intercept and slope coefficients. x_1 consists of explanatory variables including methodology, site and context characteristics, and ε_1 is the error term with mean zero and variance σ^2_{ε}.

Robust standard errors are reported to control for potential heteroskedasticity and autocorrelated estimates. A random-effects model (REM) is also estimated, as suggested by Rosenberger and Loomis (2000) and Nelson and Kennedy (2009), to account for intra-study dependency among value estimates following the structure of a specific disturbance component μ_{1j} and error component ε_1, as shown in equation (2).

$$y_{1j} = \alpha + \beta' x_{1j+}\ \varepsilon_{1+}\ \mu_{1j} \qquad (2)$$

In sum, four models were estimated, as shown in Table 3,including OLS and REM specifications for all the included studies (the “all” data model) and the CVM subset of studies. In the case of REM, the estimates were clustered by study, implying that there is a lack of independence for WTP estimates within a study; in other words, the WTP estimates from one study are likely to be correlated with one another, rather than with those between studies.

Additionally, Breusch–Pagan Lagrangemultiplier tests were performed to check whether the random-effects models are favoured over the OLS models for the all and CVM data sets. The REM model is favoured over the OLS model for the all data ($\alpha = 0.05$) and CVM data ($\alpha = 0.10$).The variance inflation factor (VIF) test suggests no multicollinearity (i.e., the VIFs are all below 10).

Table 3: Meta-analysis regression results: ordinary least squares (OLS)and random-effects model (REM)(standard errors in parenthesis)

Variable	**OLS, weight, robust ALL**	**OLS, weight, robust CVM**	**REM, weight, robust ALL**	**REM, weight, robust CVM**
Sample size	0.2404	0.4613	-0.2245	0.2026
	(0.5584)	(0.4919)	(0.5308)	(0.4057)
Personal mode	2.2467***	3.1877***	2.2540***	2.2557***
	(0.6958)	(0.5763)	(0.7834)	(0.5109)
Referendum	-6.4568***	-7.0461***	-5.6273***	-7.9093***
	(1.5359)	(1.6914)	(1.5058)	(1.3672)
Picture	-0.2121	0.3147	-0.6324	0.903
	(0.5159)	(0.7463)	(0.7847)	(1.0291)
Substitute	0.3847	0.1493	0.8171	-0.2107
	(0.4762)	(0.4813)	(1.0002)	(0.7754)
Cheap script	-0.4623	-0.8186	0.1266	-0.3872
	(0.7763)	(0.613)	(0.7643)	(0.6814)
Certainty	0.1781	-0.1698	-0.5822	-0.4069
	(0.6606)	(0.5909)	(0.7497)	(0.404)
CVM	-1.9754*		-1.3806	
	(0.9956)		(0.8739)	
CE	-4.4447***		-5.0121***	
	(1.4068)		(1.431)	
Mail	1.6018**	1.8380**	1.2407	1.2299**
	(0.729)	0.7521	(0.78)	(0.5459)
On-site	-1.4547**	-1.1003	-1.7304*	0.1779
	(0.5727)	0.7054	(0.8838)	(0.4538)
Open ended	-2.9177*	-1.3479	-5.3667***	-1.9185
	(1.5916)	0.9571	(1.9853)	(1.1994)
Payment card	-1.9125	-0.6362	-1.8233*	-0.9994***
	(1.2511)	0.5806	(1.0839)	(0.2954)
Entrance fee	-2.2453**	-1.9068***	-2.4981**	-0.1727
	(0.8736)	(0.6586)	(0.9756)	(0.6752)
Translation	-0.5554	-0.6373	0.1397	-0.3805
	(0.7512)	(0.5039)	(0.7872)	(0.5152)
International user	1.0452**	0.3798	0.9124**	0.2928

	(0.5109)	(0.3325)	(0.4297)	(0.2341)
Year	-0.0298	0.0176	0.0133	0.0202
	(0.0782)	(0.1163)	(0.0981)	(0.1285)
Publication	0.6295	0.3049	0.1105	0.8489
	(0.444)	(0.5542)	(0.7207)	(0.712)
Natural reef	1.2229**	1.1662**	0.5753*	0.7165***
	(0.5058)	(0.4932)	(0.3219)	(0.2614)
MPA	0.0603	0.3635	-0.1574	-0.0208
	(0.5914)	(0.6943)	(0.5598)	(0.599)
GDP per capita	0.1143***	0.1314***	0.1239***	0.1642***
	(0.0213)	(0.0312)	(0.0284)	(0.0274)
Coral/MPA	0.0276***	0.0243***	0.0374***	0.0245***
	(0.0078)	(0.0073)	(0.0116)	(0.0075)
CBD	-0.2998	0.8983	-1.1749	0.8941
	(0.8466)	(0.7918)	(1.1526)	(0.7872)
Fee_closed	2.3448**	2.5344**	3.3747***	0.7843
	(1.099)	(1.1964)	(1.0068)	(0.7452)
Sample_closed	2.7951	3.8639**	0.9789	5.6252***
	(1.7854)	(1.7127)	(2.1443)	(1.8308)
Constant	2.0255	-2.9768***	3.5956	-3.1995**
	(1.9586)	(1.0684)	(2.3107)	(1.5078)
N	131	89	131	89
No. of groups	46	32	46	32
R^2(OLS)	0.76	0.81		
Within errors R^2			0.60	0.67
Between errors R^2			0.71	0.75
Overall R^2			0.69	0.7
VIF	8.26	9.91		

Robust standard errors in parenthesis and significance indicated as follows: * p<0.1; ** p<0.05; *** p<0.01.

CE and CVM studies generally provide lower WTP estimates than TCM studies (the omitted category). This is consistent with previous meta-analyses, such as those by Johnston et al. (2006), Brander et al. (2007) and Londoño and Johnston (2012), who reported similar findings. In general, studies with sample sizes greater than 100 have higher WTP estimates as well as greater statistical efficiency. Also, as suggested by the NOAA, the sample size should vary depending on the type of question format used. Interacting the sample size with close-ended formats (Sample_closed) was positive for both CVM models, suggesting that larger sample size studies using closed-ended questions result in larger WTP estimates.This is confirmed by studies such as that of Jordan and Elnagheeb (1994), who found that referendum questions or closed-ended questions weremore highly sensitive to a large sample size than payment card or open-ended formats.

The elicitation format, namely personal interviews and mail surveys, are found to be positively associated with higher WTP. Additionally, for the various elicitation modes, the closed-ended question negatively influences the WTPas well as the payment card and open-ended formats. The effects, however, were slightly lower in closed-ended than payment card and/or open-ended ones when the base case was a combined mode.

The influence of the entrance fee on WTP values is negative, implying a direct cost to access such sites. A closer inspection of this variable and the interaction with closed-ended questions (Fee_closed) revealed a positive influence on WTP. This illustrates the importance of selecting a suitable elicitation mode, in this case close-ended as opposed to open-ended questions. As recommended by Green et al. (1998), the use of referendum questions should be carefully examined with the influence of the payment vehicle, as in this case entrance fees as well as framing, agenda setting and experimental bid designs affect WTP.

Non-local users' WTP is higher than local users' WTP, as expected; those who are out of the area are willing and able to pay more for recreation activities than their local counterparts. Similarly, a natural reef has a positive effect on WTP. A plausible explanation can be attached to the natural habitat, particularly for non-local users who are not close to the site. With regard to on-site responses, it is likely that close experience and/or contact with the good/service and potentially lower budgets may constrain WTP estimates.

Augmenting meta-data with secondary measures is a recommended and fruitful way to incorporate socio-economic differences across sites: information that is typically under-reported in primary studies (Johnston and Rosenberger 2010). For example, GDP per capita may serve as a proxy for local income. The coefficient GDP per capita is positive and highly significant, meaning that the higher the income for the sample, the more they are willing and able to pay for the recreation service or goods. Similarly, the reef coverage to MPA area ratio is positive, suggesting that the more the reef is closely defined as an MPA, the greater value it takes on.

Importantly, there was no presence of hypothetical bias as previously thought for substitutes, certainty follow-ups, cheap talk script and related interaction terms with the sample and/or elicitation formats/mode. Additionally, no effects on WTP estimates were found in the use of photos, publication type, questionnaire translation, data year and CBD. Even though some hypothetical bias and other bias treatments were not statistically significant in this body of applied valuation research, this does not mitigate the importance of bias, as noted in individual studies. This is especially true when attempting to use a body of results for benefit transfer purposes. The implication of methodological effects/hypothetical biases on WTP estimates suggests that the benefit transfer functions should factor into account the nature of diverse economies, particularly when the costs and expertise needed to ameliorate such effects/biases vary across the globe. In other words, developing countries differ from developed countries in their socio-economic and political structures, making the NOAA recommendations for CVM relatively difficult and costly to implement in the former compared with the latter.

Despite the challenges of carrying out CVM in comparison with other stated preference approaches, CVM remains popular with the majority, having been applied in developed countries such as the USA, Canada and Europe. Relatively few studies have focused on developing countries. Freeman (1986) remarked that the high cost for CVM is a function of accuracy, which depends heavily on the sample size, but also attempts to reduce biases and errors. It is possible that high-income economies of coral-endowed countries are likely to reduce biases/effects due to the availability of funds for an increased sample size as well as better expertise in designing CV studies than their counterparts.

One way to examine this hypothesis is the interaction of GDP per capita with some common CVM biases. In this case, three interaction variables with GDP per capita were estimated: referendum/closed question (GDP_closed),a payment card (GDP_payment) and sample size with referendum questions (GDP_sampleclosed). All these interaction terms were significant and used in the final estimations, as shown in Table4. The results of these interaction terms with GDP per capita suggest that the increased income effects and use of a payment card as well as an increased sample size with closed-ended questions may positively affectWTP. This confirms that the amelioration of biases as identified by the NOAA guidelines is associated with higher GDP per capita, as in the case of an increased sample size with closed-ended questions and/or a payment card. However, the unexpected result from this inquiry is that when GDP per capita is interacted with referendum questions and/or closed-ended questions, the result is negative, although this negative influence is consistent with previous models in Table 3. Other biases/effects, such as substitute, picture, sample size, cheap script, etc.,were insignificant in the expanded model.

Both the OLS CVM and the REM CVM model are reported in Table 4; however, the Breusch–Pagan Lagrange multiplier (BPLM) test rejects the latter (REM) in favour of the former model (OLS) when the GDP per capita and the interactions terms are added to the model. Consequently,the OLS CVM was selected as the best model, as the overall R^2 of this model was slightly improved over the previous one (Table 3).

In sum, for studies restricted by budget constraints, in particular in low-income economies, reduced sample sizes may downwardly influence WTP.As Whittington (2002) pointed out, low-income countriesare far from being a high-quality option at a low cost when taking into account sample size. That is to say, low-income countries are budget-constrained, hence limiting their ability to achieve higher sample sizes. Arkesteijn and Oerelemans (2005) and Han et al. (2008) asserted that this issue may be more prevalent in developing economies, where limited funding coupled with the high cost of surveys constrains sample sizes.

Table 4: CVM models with interaction terms

(standard errors in parenthesis)

Variable	**OLS_CVM**	**REM_CVM**
Sample size	0.2312	0.2694
	(0.4468)	(0.3475)
Personal mode	2.3775***	1.0974
	(0.7915)	(0.7357)
Referendum	-1.2906	-0.6449
	(1.1384)	(1.4636)
Picture	0.1477	0.9548
	(0.6280)	(1.0210)
Substitute	-0.0594	-0.8719
	(0.3670)	(0.8089)
Cheap script	-0.5209	-0.1831
	(0.5230)	(0.5746)
Certainty	-0.3804	-0.3381
	(0.5579)	(0.4253)
Mail	1.0709	0.0640
	(0.9423)	(0.7570)
On-site	-1.3962**	0.0636
	(0.5877)	(0.4841)
Open-ended	-0.7122	-1.1255
	(0.9491)	(1.3254)
Payment card	-0.7029	-1.3410***
	(0.5615)	(0.2274)
Entrance fee	-1.6493***	-0.0445
	(0.5873)	(0.6414)
Translation	-0.4908	-0.3924
	(0.3918)	(0.5032)
International user	0.3226	0.2841
	(0.2795)	(0.2223)
Year	0.0522	0.1168
	(0.0894)	(0.1078)
Publication	-0.0162	0.1727
	(0.4784)	(0.6085)
Natural reef	0.6825***	0.6666***
	(0.1556)	(0.2518)
MPA	-0.1960	-0.4199
	(0.5831)	(0.4756)
GDP per capita	0.1256***	0.1160***
	(0.0308)	(0.0392)
Coral/MPA	0.0258***	0.0216**

	(0.0077)	(0.0093)
CBD	0.4108	0.8308
	(0.6195)	(0.6935)
Fee_closed	1.8534*	0.5796
	(0.9764)	(0.6998)
GDP_payment	0.1265**	0.1726**
	(0.0491)	(0.0811)
GDP_closed	-2.4891**	-3.6895***
	(0.9502)	(1.0466)
GDP_sampleclosed	2.4435**	3.6645***
	(0.9436)	(1.0592)
Constant	-2.0056***	-3.1727**
	(0.7123)	(1.2522)
N	89	89
No. of groups	32	32
R^2(OLS)	0.87	
Within errors R^2		0.68
Between errors R^2		0.84
Overall R^2		0.76
VIF	7.59	

Robust standard errors in parenthesis and significance indicated as follows: * p<0.1; ** p<0.05; *** p<0.01.

Coral reef valuation has an important global application; therefore, the distinction of biases/effects with income, in this case GDP, may suggest that the benefit transfer from study to global policy sites should be calibrated for income. Subsequently, study and policy sites are allocated to one of the four economy categories identified by the World Bank income classification scheme: low, lower-middle, upper-middle and high. Next, values are forecast from the meta-analysis function (the 35 coral-endowed countries and territories divided into the four economy categories) for 49countries and territories (policy sites). The differentiated mean valueby income class is then used in the international benefit transfer, for which the predictionswere estimated for policy sites, except for the coral/MPA ratio and GDP per capita, for which the respective values from the policy sites were used.

The forecasted values were estimated according to the Rosenberger and Loomis (2001) approach, in which the base-case scenario uses mean values for the included variables, and bias and no bias scenarios either turn on (1) or off (0) bias effects in the BT exercise. The results are reported in Table 5 using the OLS CVM model as the preferred model based on statistical tests.

Table 5: Meta-analysis benefit function transfer scenarios for varied economies

	Mean recreation BT (US$)	Proportion of BT to GDP per capita		Mean recreation BT (US$)	Proportion of BT to GDP per capita
a) Low (mean GDP per capita=850)			b) Lower-middle (mean GDP per capita=2,000)		
Base case	101	0.12	Base case	587	0.29
Bias	369	0.43	Bias	2,051	1.03
No bias	45	0.05	No bias	252	0.13
Bias minus base (a)	268	0.32	Bias minus base (a)	1,464	0.73
Bias minus no bias (b)	324	0.38	Bias minus no bias (b)	1,799	0.90
Ratio of (a)/(b)		0.84	Ratio of (a)/(b)		0.81
c) Upper-middle (mean GDP per capita=6,750)			d) High (mean GDP per capita=17,000)		
Base case	3,248	0.48	Base case	10,653	0.63
Bias	5,100	0.76	Bias	11,597	0.68
No bias	798	0.12	No bias	1,427	0.08
Bias minus base (a)	1,852	0.27	Bias minus base (a)	944	0.06
Bias minus no bias (b)	4,302	0.64	Bias minus no bias (b)	10,170	0.60
Ratio of (a)/(b)		0.42	Ratio of (a)/(b)		0.10

Notes: Low countries (Madagascar, Mozambique and Tokelau), lower-middle (Bangladesh, Cambodia, Cameroon, India, Kiribati, Marshall Islands, Nicaragua, Pakistan, Papua Guinea and Solomon Islands), upper-middle (Belize, Brazil, China, Colombia, Costa Rica, Domenica, Dominican Republic, Ecuador, Honduras, Indonesia, Jordan, Mauritius, Palau, Panama, South Africa, Sri Lanka, St. Vincent and the Grenadines, Tonga, Turks and Caicos Islands, Tuvalu and Venezuela) and high (Anguilla, Antigua and Barbuda, Oman, Puerto Rico, Saudi Arabia, Trinidad and Tobago and Virgin Islands, United States).[4]

Significantly, and irrespective of the economy type, the base case, bias and no bias treatments increase with income class, though this is more pronounced for the bias and no bias cases for high economies. Nevertheless, the base case effect is relatively inflated to the no bias scenarios, for nearly all the economies. In the case of the high-income class, the difference between bias scenario and base case is relatively lower than upper-middle, lower-middle and low economies, implying that this category is able to "cushion" the biases better than other economies. Significantly, the percentage of “cushioning” is lower at 6% for the high-income class than the upper-middle (27%), lower-middle (73%) and low economies (32%). Similarly, with respect to the differences between the bias and no bias cases over GDP per capita, the trend is the same except that the low economies report a lower estimate at 38%.

As shown in Table 5, the ratio of BT estimates between the bias minus base and the bias minus no bias treatments suggests that low and low-middle economies are unfavourable with high ratios of at least 0.8. In other words, these groups have slight variations for all the treatments, suggesting that they are more affected by biases. On the one hand, low economies have the potential to manage biases better when selecting an elicitation mode and format that are favourable to their surroundings (i.e., personal elicitation). However, on the other hand, they are faced with limited funds when larger sample sizes are required to lower biases and/or errors. One plausible avenue to overcome such a challenge is to foster co-operation with high economies in research, though this requires regional and global support from funding institutions as well as valuation panels/experts. Indeed, such an effort can “perfect” the CVM approach in a rigorous environment, increase the number of valuation studies and report less biased estimates.

5.Conclusion

Numerous meta-analyses have been applied to environmental economics applied research, including specific testing for hypothetical bias tests. This paper uses meta-analysis to evaluate the effects of the NOAA valuation guidelines and hypothetical bias in a specific body of research—coral reef valuation. The empirical results of this study illustrate that, to some extent, the influence of biases and/or effects as recommended by the NOAA’s directives are relevant and significant in influencing estimates of WTP and, by extension, benefit transfer.

Nonetheless, the popularity of CVM studies for lower-income economies, despite the high cost as well as sampling issues, suggests that some methodological issues may require more attention in these economies as well as in high economies. The overall policy implication of these estimates illustrates that biases in study and policy sites may vary according to income effects, which in turn affect the “cushioning” or absorption of biases depending on the sample size, elicitation format and mode. Consequently, this can affect policy values related to conservation and management efforts for a multi-functional ecosystem such as a coral reef. However, the interpretation of these effects and/or biases should be treated with caution, particularly when the WTP values are distinguished by methodological characteristics and the influence of income, the number of studies and the spatial area of natural resources.

Furthermore, the biases/effects connected to valuation methods are within the control of valuation designers rather than those held by respondents. Hence, it is imperative for valuation designers to recognize these biases and/or effects and to explore a level playing field to mitigate them in order to achieve survey efficiency. The key lessons learned from this study suggest that recognizing such biases as pointed out in the NOAA guidelines is pertinent. A recommendation for CVM practitioners and/or researchers is to provide detailed information in their primary studies, particularly related to substitutes as well cross-tabulation, if any, as suggested by the guidelines. Moreover, what is needed is consistency in publishing values; in other words, a proposed uniform reporting template identifying the key NOAA guidelines for primary CVM studies is vital for future MRA. Importantly, CVM studies designed to estimate the environmental values pertaining to natural resources need to define the quality of such an ecosystem/biodiversity.

Another critical area for future research for the purpose of BT transfer for corals is the profile of marine protected areas (MPAs). Nearly 50% of the values reported in this database were obtained from MPA sites. Moreover, the reef coverage to MPA ratio was highly significant and positively influenced the WTP. However, what is lacking in most primary studies is further details about MPA, such as the management type, year of establishment and fee type. Such information can provide further policy implications considering multiple users of the coral sites and this in turn may affect the pricing strategies, type of institution to manage (e.g., community, non-government, government) and/or restriction or prohibition of other activities, such as fishing, anchoring or vessels in the area. In sum, this exercise has demonstrated that the NOAA guidelines in CVM studies go hand in hand with income and as such transferring WTP values from study to policy site requires calibration by considering all the plausible effects found.

ACKNOWLEDGEMENT

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no.PIEF-GA-2009-253724, project "Risks of global warming: the case of coral reef ecosystems in developing countries (BIOCORE)".

NOTES

[1] Londoño and Johnston (2012) also evaluated travel cost method estimates; however, the primary focus of their paper was on the contingent valuation literature.

[2]The complete citation details of these published/unpublished studies are available on request from the corresponding author.

[3]Such problems include lexical equivalence (equivalence in the same words), conceptual equivalence (equivalence in meaning), equivalence in measurement (equivalence measuring from one site or culture to another) and equivalence of response in a cross-cultural survey.

[4]Closer examination using a scatter plot of GDP per capita and coral/MPA area illustrated that some countries were outliers, hence their exclusion from the policy sites: Somalia, Djibouti, Sudan, Jamaica, St. Lucia, Bahrain, Brunei Darussalam, the Cayman Islands, French Polynesia, Japan, Qatar, Singapore, St. Kitts and Nevis, Vanuatu, the United Arab Emirates andthe Virgin Islands, British. Similarly, the transfer estimates from study site excluded the United States.

APPENDIX

Selected valuation studies included in the meta-analysis

	Author(s)	**Year**	**TCM**	**CE**	**CVM**	**OTH**	**Region**
1	Ahmad and Hanley	2009	0	0	x	0	South-East Asia
2	Ahmed et al.	2007	x	0	x	0	Oceania, Mela, Micro
3	Andersson	2007	0	0	x	0	Africa, East
4	Arin and Kramer	2002	0	0	x	0	Oceania, Mela, Micro
5	Barr and Mourato	2009	0	0	x	0	Central America
6	Bell et al.	1998	0	0	x	0	N. America, US
7	Beukering et al.	2006	0	x	0	0	Oceania, Mela, Micro
8	Beukering et al.	2000	0	x	0	0	Caribbean and small islands
9	Bhat	2003	x	0	0	0	N. America, US
10	Carr and Mendelsohn	2003	x	0	0	0	Australia
11	Casey et al.	2010	0	0	x	0	Central America
12	Cesar	2003	x	0	x	0	Europe, South
13	Christiernsson	2003	x	0	0	0	South-East Asia
14	Dharmaratne et al.	2000	0	0	x	0	Caribbean and small islands
15	Dixon et al.	1993	0	0	x	0	Caribbean and small

							islands
16	Edwards	2009	0	0	x	0	Caribbean and small islands
17	Hushak et al.	1999	x	0	0	0	N. America, US
18	Johns et al.	2004	0	0	x	0	N. America, US
19	Kragt et al.	2009	0	0	0	x	Australia
20	Leeworthy and Bowker	1997	x	0	0	0	N. America, US
21	Leeworthy et al.	2001	0	0	x	0	N. America, US
22	Lindsey and holmes	2002	0	0	x	0	South-East Asia
23	Mathieu et al.	2003	0	0	x	0	Africa, East
24	McCartney	2009	0	x	0	0	Australia
25	Milon	1998	x	0	0	0	N. America, US
26	Mohamed et al.	2001	0	0	x	0	Oceania, Mela, Micro
27	Mohamed	2007	0	0	x	0	South Asia
28	Morgan et al.	2010	x	0	0	0	N. America, US
29	Ngazy et al.	2004	0	0	x	0	Africa, East
30	O’Garra	2009	0	0	x	0	Caribbean and small islands
31	Oh et al.	2008	0	0	x	0	N. America, US
32	Park et al.	2002	x	0	x	0	N. America, US
33	Parsons and Thur	2008	0	x	0	0	Caribbean and small islands
34	Pendelton	1995	x	0	0	0	Caribbean and small islands
35	Pham and Son	2001	x	0	x	0	South-East Asia
36	Ransom and Mangi	2010	0	0	x	0	Africa, East
37	Rolfe and Windle	2010	0	x	0	0	Australia
38	Rudd et al.	2001	0	x	0	0	Caribbean and small islands
39	Asafu-Adjaye and Tapsuwan	2008	0	0	x	0	South-East Asia
40	Samonte-Tan et al.	2007	0	0	0	x	Oceania, Mela, Micro
41	Seenprachawong	2003	0	0	x	0	South-East Asia
42	Sorice et al.	2007	0	x	0	0	N. America, US
43	Spash	2000	0	0	x	0	Caribbean and small islands
44	Spurgeon et al.	2004	0	0	x	0	Caribbean and small islands
45	Subade	2005	0	0	x	0	Oceania, Mela, Micro
46	Svensson et al.	2008	0	0	x	0	South-East Asia
47	Talaat et al.	2009	0	x	x	0	Europe, South
48	Thur	2010	0	0	x	0	Caribbean and small islands
49	Uyarra et al.	2010	0	0	x	0	Caribbean and small islands
50	White	2008	0	0	x	0	N. America, US
51	Wielgus et al.	2003	0	x	0	0	Europe, South
52	Yacob et al.	2009	0	0	x	0	South-East Asia
53	Yeo	2004	0	0	x	0	South-East Asia

REFERENCES

Aadlund D, Caplan AJ (2006)Cheap talk reconsidered: new evidence from CVM.J Econ Behav Organ 60: 562–578.

Akter S, Bennett J, Akhter S (2008) Preference uncertainty in contingent valuation. Ecol Econ 67: 345–351.

Arkesteijn K, Oerelemans L (2005)The early adoption of green power by Dutch households. An empirical exploration of factors influencing the early adoption of green electricity for domestic purposes.Energ Policy 33: 183–196.

Arrow K, Solow R, Portney PR, Leamer EE, Radner R, Schuman H (1993) Report of the NOAA panel on contingent valuation. NOAA. Available from: www.darp.noaa.gov/library/pdf/cvblue.pdf.Cited 2 April 2011.

Bateman I, Carson R, Day B, et al (2002) A manual: economic valuation with stated preference techniques. Edward Elgar Publishing, Massachusetts, USA.

Boyle K, Poe G, Bergstrom J (1994) What do we know about groundwater values? Preliminary implications from a meta analysis of contingent valuation studies.Am J Agr Econ 76: 1055–1061.

Brander LM, van Beukering P, Cesar HSJ (2007)The recreational value of coral reefs: a meta-analysis. Ecol Econ 63(1): 209–218.

Bulmer M, Warwick DP (1993)Social research in developing countries: surveys and census in the third world. University College London (UCL), London, UK.

Carson RT, Flores NE, Meade NF (2001) Contingent valuation: controversies and evidence. Environ Resour Econ 19:173–210.

Carson RT, Groves T (2007) Incentive and informational properties of preference questions. Environ Resour Econ 37:181–210.

Champ PA, Bishop RC, Brown TC, McCollum DW (1997) Using donation mechanisms to value non-use benefits from public goods. J Environ Econ Manag 33: 151–162.

Cherry T, Shogren, J (2007) Rationality crossovers.J Econ Psychol 28: 261–77.

Convention on Biodiversity Diversity, List of parties. Available from: http://www.cbd.int/information/parties.shtml.Cited 1 May 2011.

Cummings RG, Taylor LO(1999) Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. Am Econ Rev 89: 649–665.

Freeman AM (1986) Assessing the state of arts of the contingent valuation method of valuing environmental changes. In: Cummings RG, Brookshire DS, Schulze WD (eds) Valuing environmental goods. Rowman and Littlefield Publishers, Maryland, USA.

Ghermandi A, van den Bergh JCJM, Brander LM et al. (2010) The values of natural and human-made wetlands: a meta-analysis. Water Resource Research 46: W12516.

Green D, Jacowitz KE, Kahneman D et al. (1998) Referendum contingent valuation, anchoring and willingness to pay for public goods.Resour Energy Econ 20: 85–116.

Han SY, Yoo SH, Kwak SJ (2008)Valuing multiple environmental impacts of large dam construction in Korea: a choice experiment study. Environ Impact Asses 28: 256–266.

Hanemann WM (1994) Valuing the environment through contingent valuation. J Econ Perspect 8(4):19–43.

Hanley N, Spash CL (1993) Cost-benefit analysis and the environment. Edward Elgar Publishing, Vermont, USA.

Intergovernmental Panel on Climate Change (IPCC) (1995) The science of climate change. In:, Houghton JT, Meira LG, Callender BA, Harris N, Kattenberg A, Maskell K(eds). Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK.

International Monetary Fund (IMF) World economic outlook database, April 2011. Available from: http://www.imf.org/external/data.htm.Cited 1 June 2011.

Johnston RJ, Besedin EY, Ranson MH (2006) Characterizing the effects of valuation methodology in function-based benefits transfer. Ecol Econ 60(2): 407–419.

Johnston RJ, Rosenberger RS (2010) Methods, trends and controversies in contemporary benefit transfer. J Econ Surv 24(3): 479–510.

Jordan, LJ, Elnagheeb, AH (1994) Differences in contingent valuation estimates from referendum and checklist questions. J Agr Resour Econ 19(1): 115–128.

List JA, Gallett CA (2001) What experimental protocol influence disparities between actual and hypothetical stated values? Environ Resour Econ 20: 241–254.

List JA, Paramita S, Taylor MH (2006) Using choice experiments to value non-market goods and services: evidence from field experiments. Adv Econ Anal Pol 6(2-Advances), Article 2

Londoño LM, Johnston RJ (2012) Enhancing the reliability of benefit transfer over heterogeneous sites: A meta-analysis of international coral reef values. Ecol Econ 78: 80-89

Loomis JB, White DS (1996) Economic benefits of rare and endangered species: summary and meta-analysis. Ecol Econ 18: 197–206.

Mitchell R, Carson R (1989) Using surveys to value public goods.Resources for the Future, Washington DC, USA.

Murphy JJ, Geoffrey AP, Stevens TH et al. (2005a) A meta-analysis of hypothetical bias in contingent valuation. Environ Resour Econ 30(3): 313–325.

Murphy JJ, Stevens TH, Weatherhead D (2005b) Is cheap talk effective at eliminating hypothetical bias in a provision point mechanism? Environ Resour Econ 30: 327–343.

Nelson JP, Kennedy PE (2009)The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment.Environ Resour Econ 42(3): 345–377.

Noonan DS (2003) Contingent valuation and cultural resources: a meta-analytic review of literature.J Cult Econ 27(3–4): 159–176.

Ready R, Whitehead J, Blomquist G (1995) Contingent valuation when respondents are ambivalent. J Environ Econ Manag 29: 181–197.

Rosenberger RS, Johnston RJ (2009) Selection effects in meta-analysis and benefit transfer: avoiding unintended consequences. Land Econ 85: 410–428.

Rosenberger RS, Loomis JB (2000) Panel stratification in meta-analysis of environmental and natural resource economic studies. J Agr Resour Econ 32(3): 459–470.

Rosenberger RS,Loomis JB (2001)Benefit transfer of outdoor recreation use values: a technical document supporting the Forest Service Strategic Plan (2000 revision).U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, RMRS-GTR-72,Fort Collins, CO.

Rosenberger RS, Stanley TD (2006) Measurement, generalization and publication: sources of error in benefit transfers and their management. Ecol Econ 60:372–378.

Stanley TD (2001) Wheat from chaff: meta-analysis as quantitative literature review. J Econ Perspect 15:131–150.

Woodward RT, Wui YS (2001)The economic value of wetland services: a meta-analysis. Ecol Econ 37:257–270.

World Bank (2010) World Development Indicator (WDI), World Bank, Washington DC. Available from: http://data.worldbank.org/indicator.Cited 1 April 2011.

World Databases on Protected Areas (2009) World database on protected areas annual report. Available from: http://www.wdpa.org/AnnualRelease.aspx.Cited 10 June 2011.

Whittington D (2002) Improving the performance of contingent valuation in developing countries.Environ Resour Econ 22: 323–367.