

Valuation of species and nature conservation in Asia and Oceania: a meta-analysis

Henrik Lindhjem · Tran Huu Tuan

Received: 18 January 2011 / Accepted: 25 July 2011 / Published online: 20 August 2011
© Springer 2011

Abstract We conduct a meta-analysis (MA) of around 100 studies valuing species and nature conservation in Asia and Oceania, using both revealed and stated preferences methods. Dividing our dataset into two levels of heterogeneity in terms of good characteristics (species vs. nature conservation more generally) and valuation methods, we show that the degree of regularity and conformity with theory and empirical expectations is higher for the more homogenous dataset of contingent valuation of species. For example, we find that willingness to pay (WTP) for preservation of mammals tends to be higher than other species and that WTP for species preservation increases with income (elasticity below one). For the full dataset we find that marine habitats are valued significantly higher than other habitat types in the region. Despite some encouraging results, more research is required to answer the question of how homogenous is homogenous enough in MA, especially when moving towards using MA for benefit transfer and policy use.

Keywords Asia · Biodiversity · Meta-analysis · Oceania · Valuation

JEL Classification Q26 · Q51 · Q57 · H41

Electronic supplementary material The online version of this article (doi:[10.1007/s10018-011-0019-x](https://doi.org/10.1007/s10018-011-0019-x)) contains supplementary material, which is available to authorized users.

H. Lindhjem (✉)
Norwegian Institute for Nature Research (NINA), Gaustadalleen 21, 0349 Oslo, Norway
e-mail: henrik.lindhjem@nina.no

T. H. Tuan
College of Economics, Hue University, 100 Phung Hung Street, Hue, Vietnam

1 Introduction

According to the Millennium Ecosystem Assessment, more than 60% of the world's ecosystems are being degraded or used unsustainably (Millennium Ecosystem Assessment 2005). The pressure on nature is among the highest in the many rapidly growing economies of Asia and Oceania. The (neoclassical) economist's prescription to stemming this trend is to value changes in the provision of environmental goods in monetary terms and create mechanisms to internalise their values in the billions of everyday decisions of consumers, producers and government officials. In response to this challenge, an enormous amount of primary valuation research has been produced using stated and revealed preference methods in Asia and elsewhere. However, paraphrasing Glass et al. (1981: p11),¹ results of much of this work "are strewn among the scree of a hundred journals and lies in the unsightly rubble of a million dissertations". This valuation research could be much better utilised to demonstrate the social return to nature conservation, a key area where environmental economists need to do more in the future, as pointed out by the late David Pearce (2005) and more recently by the international study "The Economics of Ecosystems and Biodiversity".²

For a range of environmental goods meta-analysis (MA) techniques have been used to synthesise valuation research, test hypotheses and facilitate the transfer of existing welfare estimates to new, unstudied policy sites ("benefit transfer"—BT) for use, e.g. in cost–benefit analysis (Smith and Pattanayak 2002; Navrud and Ready 2007). Responding to Pearce's challenge, this is to our knowledge the first MA that reviews and takes stock of the literature on environmental valuation of a complex and somewhat heterogeneous good: (changes in) conservation of habitat, biodiversity and species, in a geographical region where such research is rapidly growing: Asia and Oceania.

We attempt to answer the following two research questions: (1) To what extent do welfare estimates for this complex good conform with theoretically and empirically derived expectations regarding the good characteristics, valuation methods, study quality, socio-economic and other variables?; and (2) how sensitive are the meta-regression results to; (a) the "scope of the MA", i.e. the level of heterogeneity of the good valued and the valuation methods used; and (b) the choice of meta-regression models? The first question investigates whether the welfare estimates display the degree of validity and regularity more typically found for less complex environmental goods with higher share of use values and offers a first check of the potential for using such data for BT applications (Johnston et al. 2005; Lindhjem 2007).

The second question contributes to our understanding and refinement of MA methodology in environmental economics, where the meta-analyst typically is left to make a number of choices potentially introducing various subjective biases (Hoehn 2006). An important choice both for the robustness of MA models and their suitability for use in BT applications relates to the scope of the MA, i.e. the trade-off

¹ Originally quoted in Stanley and Jarrel (2005).

² <http://www.teebweb.org/>.

between the number of observations and the acceptable level of heterogeneity of the data, as pointed out by, e.g. Engel (2002) and Nelson and Kennedy (2009) (question 2a above). Another important consideration is the choice among meta-regression models, for example which covariates to include (question 2b).³

Previous MA studies have primarily analysed the values of more homogenous types of environmental goods (e.g. water and air quality, recreation days) often within the same country (Desvousges et al. 1998; Rosenberger and Loomis 2000a; Van Houtven et al. 2007). However, there is a trend towards using MA to study more complex goods in international settings, though this work has generally not considered heterogeneity in the data or related methodological choices (for goods such as wetlands, coral reefs, forests, biodiversity, agricultural land preservation) (Brander et al. 2006, 2007; Lindhjem 2007; Jacobsen and Hanley 2009; Richardson and Loomis 2009; Barrio and Loureiro 2010). Before moving towards widespread use of MA results not just as a quantitative literature review but also for BT, we think it is worth more carefully considering the questions we pose here than the literature has done to date (Lindhjem and Navrud 2008).

We begin by explaining the conceptual framework underlying MA, give a descriptive overview of the valuation literature in the region and describe the data collection and coding process (next section). In Sect. 3 we explain the econometric approach to the data and in Sect. 4 the main results. To investigate the effect of MA scope, we divide our dataset into two levels of heterogeneity; species (more similar good and methods used) and nature conservation more generally (more heterogeneity in good and methods used). We then estimate a number of random-effects meta-regression models for these two main datasets using different cleaning procedures and subsets of the data investigating conformity with expectations, explanatory power and the robustness of results. Section 5 concludes and discusses potential implications for future research and the further use of MA for literature review and BT in non-market valuation.

2 Conceptual framework and data

2.1 Conceptual framework

We define “nature conservation” broadly as the protection or active management of any natural terrestrial or aquatic ecosystem, resource or amenity, Q . The economic value measure for an increase in the level of nature conservation (Q) is the change in the quantity and/or quality (QUAL) of Q , or some set of services provided by Q , and is referred to as consumers’ surplus (CS) or Willingness to pay (WTP). It should be noted that the available valuation methods used in economics differ in terms of the welfare measures that they estimate (see, e.g. Freeman 2003)—a potential source of heterogeneity in the data we will try to control for in our regressions. From the

³ An alternative approach to dealing with classical MA challenges, not pursued here, is to use Bayesian techniques (e.g. Moeltner et al. (2007)).

standard indirect utility function, the bid function for a representative individual j for this change can be given by (Bergstrom and Taylor 2006)⁴:

$$\text{WTP} = f\left(P_j, M_j, Q_j^T - Q_j^R, \text{QUAL}_j^T - \text{QUAL}_j^R, \text{SUB}_j^T - \text{SUB}_j^R, H_j\right) \quad (1)$$

where P = a price index of market goods (assumed constant), M = (individual or household) income (assumed constant), $Q^T - Q^R$ and $\text{QUAL}^T - \text{QUAL}^R$ are the changes in quantity and quality from a reference situation (R) to a target state-of-the-world (T), SUB = substitutes for Q available to individual j , H = non-income household or individual characteristics.

Further to make (1) elastic enough for use in MA, we assume, following Bergstrom and Taylor (2006) a “weak structural utility theoretic approach” in which the underlying variables in the bid function are assumed to be derivable from some unknown utility function, but that flexibility is maintained to introduce explanatory variables into the model, such as study design and different valuation methods, that do not necessarily follow from (1). This is the most common approach in MA, where the meta-analyst records welfare estimates from different studies and corresponding explanatory variables both informed by theory and empirical expectations.

In this process the empirical specification chosen for (1) needs to trade off the availability of information reported in valuation studies with the range of potentially relevant explanatory variables. For example, information about substitute sites to a national park will mostly not be reported even if important for WTP. In addition, if information *is* reported, for example about the exact change in nature conservation valued (e.g. in hectares protected), this change may not be easily comparable across sites and studies. No MA studies are free of this problem. Some try to map changes to a common unit of measurement in terms of hectares or to a water quality ladder or similar, though such simplified common units may mask differences in other dimensions of the good important to individuals (see, e.g. Lindhjem 2007). There are no easy solutions, and in our rather general case we interpret mean WTP from different studies as welfare estimates for a (small, though not marginal) change in Q and/or in one or more elements in an attribute vector of QUAL describing the quality of the nature site.⁵ We then use dummy variables to detect differences in WTP depending on the type of habitat or change valued. For example, when considering studies that value preservation of biodiversity we use variables for types of species and other characteristics of the good to capture variation in this overall value category. Before discussing the empirical specification of (1), we first describe the data used for the MA.

⁴ For simplicity and brevity we do not elaborate the details of how nature conservation may increase utility e.g. related to market goods and household production, e.g. as done by Van Houtven et al. (2007) for water quality.

⁵ The ecosystem services and functions and total economic value from nature and biodiversity conservation are discussed in depth elsewhere, and not elaborated in detail here (see e.g. Fromm 2000).

Table 1 Regional distribution of valuation studies (WTP in US\$ 2006)

Region	Mean WTP (St. dev.)	No. of obs.	No. of studies
O Australia, Micronesia, Papua New Guinea, Vanuatu; SEA	83 (212)	244	32
Cambodia, Indonesia, Laos, Malaysia, the Philippines, Singapore, Thailand, Vietnam;	254 (914)	116	23
EA China, Japan, Korea, Taiwan; SA India, Sri Lanka; SWA Iran, Israel, Pakistan	76 (108)	99	23
	206 (286)	70	11
	66 (78)	21	6
Total		550	95

2.2 Description of meta-data

Given this conceptual framework, we conducted a broad search for studies (published papers, reports, book chapters, etc.⁶) internationally available in English valuing nature conservation in the region drawn from various databases, including Environmental Valuation Reference Inventory (EVRI), ECONLIT, ISI Web of Science, Environment and Economy Programme for Southeast Asia (EEPSEA’s), etc.⁷

The first studies were conducted in Australia in the 1980s. In the rest of Asia, valuation started much later, but has grown in number substantially during the 1990s and 2000s. Based on the literature search we compiled a gross meta-dataset of 577 mean WTP estimates (i.e. observations) from 99 studies. A first crude screening of the studies was conducted by excluding the ones reporting negative mean WTP or very high or low estimates (2 standard deviations of the mean), leaving 550 estimates from 95 studies for detailed analysis. This reduces the influence of outlier estimates in regressions. The resulting distribution of studies by region, by type of habitat or service valued and valuation method used is given in Tables 1, 2, 3 below.⁸ Most of the studies are from Southeast Asia, East Asia or Oceania (mostly Australia), with a smaller number of studies from South and Southwest Asia (Table 1). Australia has the largest number of studies (22), followed by the Philippines with ten studies. Raw mean annual WTP is highest for Oceania at US\$ 254, as expected, though also high for South Asia (US\$ 206). The lowest WTP, all at around the same level, is found in Southeast Asia (US\$ 83), East Asia (US\$ 76) and Southwest Asia (US\$ 66).

The most frequently valued habitat is terrestrial (including forests, nature reserves and national parks) grouped together here for ease of exposition (Table 2). Marine and freshwater habitats (i.e. coral reefs, beaches, sea, rivers and water courses) for simplicity termed “marine habitats”, follow second.

⁶ We did not include Master theses for practical reasons (hard to find and/or to get hold of) and because many are written in the native language.

⁷ Since the Australian database ENVALUE is no longer updated, has been (partly) integrated with EVRI and includes limited study information, our main search used the EVRI database.

⁸ We do not claim to have collected an exhaustive database of all studies in Asia and Oceania, but we are confident that we cover the majority of such studies in the region until 2009. Further, it is unlikely that our search has been biased in any way.

Table 2 Distribution of valuation studies by habitat types (WTP in US\$ 2006)

Types of habitats/services	Mean WTP (St. dev.)	No. of obs.	No. of studies
Terrestrial habitats (reserves, national parks, forests)	116 (252)	176	33
Marine habitats (reefs, beaches, sea, watercourses)	80 (97)	162	27
Species (single or multiple)	105 (220)	129	16
Wetlands (wetlands, mangroves)	514 (1503)	41	8
Other habitats/services (landscapes, eco.-services)	121 (182)	37	13
Total		550	97 ^a

^a Some studies have more split samples asking different types of good, and thus the number of studies is higher than reported in Table 1

Table 3 Valuation studies by methods (WTP in US\$ 2006)

Method	Mean WTP (St. dev.)	No. of obs.	No. of studies
Contingent valuation method (CV)	124 (505)	417	77
Choice modelling/experiments (CM)	67 (41)	50	8
Travel cost method (TCM)	161 (162)	37	14
Others (market price, hedonic pricing)	269 (435)	46	5
Total		550	104 ^a

^a In some studies, there are more than one method used, and thus the number of studies is higher than reported in Table 1

Wetlands have the highest value at US\$ 514 mostly due to the market price methods often used to value such habitats (see Table 3). Studies that value named species or groups of species are categorised as “species”. Marine habitats provide the lowest value (US\$ 80) compared with other types of habitats, while terrestrial habitats (US\$ 116), species (US\$ 105) and other habitats (US\$ 121) have values that are around 40–50% higher. By far, the most frequently used method of valuation is contingent valuation (CV), with 77 studies, while the travel cost method (TCM) comes second with only 14 studies (Table 3). A small number of studies (5) use other methods, such as the hedonic pricing method (HPM) or calculate the value of wetlands and forests using the market price approach. As noted earlier, these methods frequently calculate a different welfare measure than CV, choice modelling (CM) and TCM studies and also yield higher estimates than the other methods.

Details of the individual studies (including references) are given in the supplementary Appendix.

2.3 Coding of data for meta-regression analysis

Information from the studies was coded in a spreadsheet originally containing 30 of the likely most important variables chosen from the universe of potentially interesting and relevant covariates, with between 1 and 36 observations drawn from

each study (average 5.8). The same study typically has several sub-samples varying the methods used, scope and other aspects of the good being valued. Table 4 below gives the variable names and definitions we include in the final regressions. We coded a number of additional variables that were tested in preliminary analysis. However, a choice was made to include those variables that were more commonly significant or are relevant for comparing regions and for BT.

Since there is no standardised way of reporting welfare estimates in the literature, a wide variety of units are typically used, e.g. WTP per individual or household, per unit of area,⁹ per visitor, for different time periods (e.g. per month, per visit, per year, one-time amount, etc.) and in different currencies and reporting years. To deal with this problem, we standardised the values to a common metric following standard MA practice, i.e. WTP (US\$ in 2006 prices) per household per year as a default and coded WTP per individual, WTP per month, etc., using dummies. For WTP per visit from CV or TCM studies, we calculated WTP per visit per year (if the study had information about how many trips a person would make per year, we converted to WTP per year).

Values from different years were converted to 2006 prices using GDP deflators from the World Bank World Development Indicators. Purchase power parity (PPP) corrected exchange rates were used to correct for differences in price levels between countries. This is the recommended procedure in international BT and MA (Ready and Navrud 2006). Some theoretical models predict that WTP given per household is higher than individual WTP, though empirical evidence is mixed (Lindhjem and Navrud 2009). It can also be expected that WTP given per month multiplied by 12 to convert to an annual amount is higher than WTP originally stated on an annual basis (a well-known bias).

We also included other methodological variables that are often used in MA studies: whether the study was a stated preference study (i.e. CV or CM) or other methods, whether it used personal interviews, if the CV method applied a dichotomous choice (DC) question format (i.e. the respondent says yes or no to a given bid, rather than stating max WTP), whether the CV data were analysed using non-parametric statistical methods and whether the payment vehicle was formulated as voluntary (e.g. donation) or mandatory (e.g. tax).¹⁰

Some studies find that CV yields lower WTP than revealed preference studies (e.g. Carson et al. 1996), which is also in line with results in Table 3 above. DC formats are often found to give higher mean WTP than open-ended formats (a main reason is so-called yea-saying), while non-parametric methods in stated preference such as the Turnbull estimator, typically give a conservative lower bound on WTP (see, e.g. Bateman et al. 2002). There is no clear prior for use of interviews versus other modes, though type of survey mode is known to influence results (Lindhjem and Navrud 2011a, b). There is also mixed a priori expectations regarding the

⁹ Studies that reported results with per unit of an area were excluded, as the total size typically was not given.

¹⁰ See the tables in the supplementary Appendix for classifications of the studies along some of these main dimensions.

Table 4 Definition of meta-analysis variables and descriptive statistics

Variables	Description	Mean (SD) ^a
Dependent variable		
WTP 2006	WTP in 2006 prices (US\$)	133 (461)
Methodological variables		
SP	Binary: 1 if stated preference, 0 if otherwise	0.84 (0.35)
DC	Binary: 1 if SP using dichotomous choice, 0 if otherwise	0.51 (0.50)
TCM	Binary: 1 if travel cost method, 0 if otherwise	0.07 (0.25)
Hholdpay	Binary: 1 if household's WTP, 0 if individual	0.67 (0.46)
Month	Binary: 1 if payment is a monthly payment, 0 if otherwise	0.35 (0.47)
Nonpara	Binary: 1 if estimate is non-parametric (Turnbull), 0 otherwise	0.07 (0.25)
Interview	Binary: 1 if it is an in-person interview, 0 otherwise	0.60 (0.48)
Mandatory	Binary: 1 if it is a mandatory payment vehicle, 0 if voluntary	0.69 (0.88)
Good characteristics variables		
Mammal	Binary: 1 if it is a mammal, 0 otherwise	0.04 (0.20)
Turtle	Binary: 1 for sea turtle, 0 otherwise	0.06 (0.24)
Species	Binary: 1 for species, 0 if other habitats/services	0.23 (0.42)
Terrestrial	Binary: 1 for terrestrial habitats, 0 if other habitats/services	0.32 (0.47)
Marine	Binary: 1 if marine habitat (beach, sea, watercourse, lake, river), 0 other habitats/services	0.29 (0.45)
Wetland	Binary: 1 for wetlands, 0 if other habitats/services	0.07 (0.26)
Nonuse	Binary: 1 for primarily non-use, 0 otherwise	0.77 (0.41)
Socio-economic variables		
Income	Continuous: Mean household income from sample (PPP adjustment, 2006)	14,318 (17,258)
GDP	Continuous: GDP 2006 from country for survey. Inserted for household income in one model.	14,524 (12,191)
Geographic characteristics (countries and regions)		
Australia	Binary: 1 if the study in Australia, 0 otherwise	0.19 (0.39)
Philippin	Binary: 1 if a study in the Philippines, 0 otherwise	0.22 (0.42)
Oceania	Binary: 1 if a study in Oceania, 0 other region	0.21 (0.40)
East	Binary: 1 if a study in East Asia, 0 other region	0.18 (0.38)
Southeast	Binary: 1 if a study in Southeast Asia, 0 otherwise	0.44 (0.48)
Southwest	Binary: 1 if a study in Southwest Asia, 0 otherwise	0.04 (0.19)
South	Binary: 1 if a study in South Asia, 0 otherwise	0.13 (0.33)
Other variables		
EEPSEA	1 if the study is funded by EEPSEA, 0 otherwise	0.39 (0.48)
Journal	1 if it is a published paper, 0 otherwise	0.47 (0.49)
Year	Continuous: from 0 (2006) to 26 (1979)	6.36 (4.07)

EEPSEA Environment and Economy Program for Southeast Asia

^a The mean (SD = standard deviation) is for overview purposes given for the whole dataset. The scope of the dataset is limited in the model runs in the next section. Further not all variables are used in all models

voluntary versus mandatory payment vehicle (e.g. Jianjun et al. 2008) (see discussion in results section below).

Further, we include a set of geographic and good characteristics variables to control for differences in welfare estimates between types of species (mammals, turtles) and habitat types, between regions and countries and between primarily non-use versus use value. Larger and more charismatic or iconized species (for example elephants or pandas) are likely to yield higher welfare estimates than non-charismatic species or biodiversity/nature conservation in general (e.g. as found in Jacobsen et al. 2008 and Richardson and Loomis 2009),¹¹ though it is uncertain a priori if our MA will be able to detect such a pattern across several studies. Studies that primarily estimate non-use values may give rise to both higher and lower value estimates than studies that cover only use value (or both use and non-use). This leaves the expected relationship of the “nonuse” variable with WTP ambiguous. There are also no strong priors regarding other habitat types or regional/country dummies, though it is expected that these dimensions may influence WTP.¹² We considered including a dummy for the season of the study (e.g. rainy vs. dry season) similar to Lindhjem (2007); however, in most cases such information was not reported.

The only socio-economic variable generally reported is income of the sample, which we include in our analysis. Around 78% of the studies report this. For those which do not, we follow common practice from other MA studies to use a proxy for income from other sources instead, i.e. we use GDP/capita for the country. It is expected that income will positively influence WTP, an often-found result in the literature for primary studies. However, in MA studies WTP is often relatively insensitive to income levels (see, e.g. Johnston et al. 2005; Jacobsen and Hanley 2009). One reason for this is the low variation in income levels in MA studies conducted within the same country or in Western countries with similar income levels. In our case we have a fairly large variation in income levels and so should expect that WTP may increase with income.

Finally, we include a proxy variable for study quality; whether a study is a published or unpublished paper (i.e. a journal article or research report/working paper). Though published studies may be expected to apply more stringent and perhaps conservative methods, it is not clear if this would result in lower WTP. There may also be publication bias with unknown influence on WTP (Rosenberger and Johnston 2009). A way to limit the potential impact of publication bias is also to include unpublished studies. To account for trends in WTP values over time that are not captured by income (or other coded variables), we include a trend variable for the year of the study (rather than publication year). Some MA studies find WTP to increase over time, reflecting, perhaps, both increased nature scarcity and “greener” preferences. Others argue that increased methodological prudence should result in lower WTP estimates in more recent studies. Since a portion of our studies is funded

¹¹ Which also tend to be reflected in actual conservation policies (see e.g. Metrick and Weitzman 1996).

¹² We also considered using population density of the country of study as a variable, for example as done by Brander et al. (2006) for wetlands. However, we think link between nature conservation and population density may be overly tenuous, and excluded this variable in our analysis.

by the same institution, Environment and Economy Program for Southeast Asia (EEPSEA), and may share similarities we have not otherwise coded, we include a dummy (“EEPSEA”) to control for that. This procedure is similar to Bateman and Jones (2003) who find indications of similarities in WTP estimates from the same authors.

3 Meta-regression model

For our meta-regressions, we divided the dataset into two primary levels of scope, according to level of homogeneity of the good and methods used: Level 1: Species; and Level 2: Biodiversity and nature conservation more generally. The species data include WTP estimates from 16 studies using CV to value the preservation of single or multiple species. These CV studies typically ask how much local/domestic populations are willing to pay for various conservation programmes for species (e.g. WTP to conserve a viable population of sea turtles).¹³

Ten of the studies are funded by EEPSEA (hence the importance of the control variable discussed above). The species valued in these studies include sea turtles (several countries), black-faced spoonbill (Macau), rhinos (Vietnam), eagles and whale shark (Philippines) and various other species such as Dugong Dugong, elephants, rhinos, dolphins and tigers (Thailand). In addition, we found six non-EEPSEA funded studies in the region using CV to value the preservation of the possum (a marsupial species native to Australia) and glider (the Mahogany Glider: a type of endangered possum), giant panda (China) and elephants (India, Sri Lanka). These 16 studies (of species) provide 124 estimates that will be used in the meta-regression analysis. Although the species are different, we consider the preservation of them as a good with many similar attributes in valuation (i.e. a larger degree of homogeneity of the good), as compared with nature and biodiversity conservation more generally. In addition, methodological heterogeneity is reduced since all the studies in this level use CV.

The second level of the data include the studies from Level 1 and all the rest of the studies that value nature conservation more generally, with different types of methods (though the majority also use CV here). This dataset includes welfare estimates for a fairly heterogeneous good, however, not more so than many other complex environmental goods in recent MA studies. Further, as almost all non-textbook goods in general (and environmental goods in particular) are heterogeneous to some degree, it is not completely clear from theory where to draw the line in practice.

All in all the Level 2 dataset contains between 67–95 studies and 390–550 estimates, depending on the cleaning procedures and the subsets of the data used in the meta-regressions. The details of the Level 1 and 2 datasets are given in Tables A1 and A2, respectively, in the supplementary appendix. We will conduct several

¹³ A small number of studies survey foreign populations, e.g. Bandara and Tisdell (2005) study OECD citizens' WTP for the preservation of the Giant Panda in China.

meta-regression models based on these two levels of data to explain variation in welfare estimates and to investigate effects of different dimensions of heterogeneity.

As most studies provide more than one WTP estimate, the data should most prudently be treated as a panel to account for the correlation between the errors of estimates from the same study¹⁴ (e.g. Nelson and Kennedy 2009). Thus, we used the procedure proposed by Rosenberger and Loomis (2000b) to check for panel structures in the data. The panel structure model, our empirical specification of Eq. (1) above, can be written as

$$WTP_{ij} = \alpha + \sum_{k=1}^n \beta_k x_{kij} + \mu_{ij} + \varepsilon_i \tag{2}$$

where WTP is the *i*th observation from the *j*th stratum (here study), and α is a constant. The variation in WTP_{ij} is to be explained by a vector of covariates $k = 1, \dots, n$, denoted by x_{kij} (as defined in Table 4), with a panel effect μ_{ij} and an error $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$.¹⁵ We also assume that μ_{ij} , ε_i and x_{kij} are uncorrelated within and across studies. A Breusch–Pagan’s Lagrange multiplier statistic test of whether panel effects are significant was conducted. The null hypothesis is that an equal effects model is correct ($H_0 : \mu_{ij} = 0$), and the alternative that a panel effects model is correct ($H_1 : \mu_{ij} \neq 0$). If the hypothesis of fixed effects in the Breusch–Pagan Lagrangian multiplier test is rejected, the random-effects model assuming heterogeneous effect sizes across studies and within models should be more efficient in estimation. We chose a double-log specification of (2), common in the MA literature, which fitted our data better than linear or other specifications. For a model with income as the only explanatory variable,¹⁶ the Breusch and Pagan Lagrange multiplier test showed that a model with equal effects was rejected, confirming the appropriateness of a panel estimation model ($\chi^2 = 274.90$, $p = 0.000$ with $N = 550$ and $j = 95$).

In order to test whether a random-effects specification (which has a panel specific error component) is outperformed by a fixed effects model (which keeps the panel specific error component constant), a Hausman χ^2 test was performed for the whole dataset. The null hypothesis is that the random effect specification is correct, i.e. the panel effects are uncorrelated with other regressors and the alternative that the fixed effect specification is correct (Zanderson and Tol 2009). The results in Table 5 show that the random-effects model (B) cannot be rejected and thus is the one we use.

We also performed the Hausman test for all the models used in this study, i.e. for different subsets of the data and different explanatory variables included and find

¹⁴ We also tested two other stratifications of the data: by survey and by author. Results (available from the authors) show that in many model specifications of the two stratifications equal effects (and random effects) cannot be rejected.

¹⁵ Standard error of WTP estimates was generally not reported, making it impossible to weigh estimates by level of precision in the meta-regressions, a procedure recommended in the MA literature (e.g. USE PA (2006)). Using the sample sizes as proxy would also loose to many observations.

¹⁶ A comprehensive test would have included other explanatory variables with different model specifications, but for sake of simplicity and brevity, we only present the model with the income variable here.

Table 5 Test of random versus fixed effects panel structure ($N = 550, j = 95$)

	b Fixed effects model	B Random-effects model	b – B	SE
Income variable ($p > \chi^2$: 0.7155)	0.0305127	-0.0494427	0.0799554	0.2193994

that a random-effects model is the best estimation approach for Level 1 and 2 of our data.

4 Meta-regression results

4.1 Level 1: Species

First, we provide results for different meta-regression models for the Level 1 (species) data. Results of six random-effects GLS regression models are reported in Table 6. Starting with Model 5, this is a model that includes all explanatory variables in Table 4 of relevance to the Level 1 data. Only one regional dummy and two species type dummies (instead of the full range of species types) are used, as estimates are thinly spread across categories.

Model 5 shows how a fully specified model including all the most relevant explanatory variables is able to deal with the heterogeneity of the data, for the more homogenous of the two datasets. Models 1–4 (and 6) are versions of Model 5 where adding different subcategories of variables illustrates changes in the explanatory power of the models. Model 1 contains methodological characteristics of the CV method only, Model 2 adds good characteristics, Model 3 adds country variables (instead of region dummy in Model 5), and Model 4 includes income and the survey year variables. A range of models was tried using combinations of variables in Table 4. Models 1–4 presented here were chosen to avoid collinearity (excluding e.g. the EEPSEA and Journal variables), to include dummies reflecting a significant share of the data (i.e. excluding region dummies for Level 1 data), to obtain best fit and to enhance comparison between models and between Level 1 and Level 2 data.

Going from Model 1 to 4, the models gradually explain more of the variation in WTP for species preservation. The methodological variables in Model 1 explain around 35% of the variation ($R^2 = 0.354$), while adding characteristics of the species explain another 15% of the variation ($R^2 = 0.500$). Adding country specifics and income and year in Models 3 and 4 helps explain other 25–31% points of the variation. Model 4, the best fitting of the models, obtains an overall R^2 of 0.81, which is very high compared with other MA studies. Model 5 obtains nearly the same level of R^2 . It is comforting for our belief in the validity of the data and for the potential use of such value estimates for BT that around half of the explained variation in the best model is due to non-study specific, observable characteristics related to the good, geographical area, year of study and income level of the population surveyed. Model 6, a version of Model 5 where all method variables are excluded, drives home the same point, with a R^2 of 61%. This model may be of

Table 6 Meta-regression models for Level I: Species studies (standard error of coefficients in parentheses)

Variable [#]	Model 1 Method variables	Model 2 +Species types	Model 3 +Country variables	Model 4 +Income and year	Model 5 All variables	Model 6 No method variables
Constant	1.149 (0.757)	2.262*** (0.834)	1.010 (1.155)	-9.949*** (3.043)	13.939*** (5.510)	-0.455 (3.861)
DC	1.462* (0.792)	0.646 (0.819)	1.191 (0.782)	1.542*** (0.537)	1.300** (0.583)	
Hholdpay	0.095 (0.789)	0.103 (0.813)	0.709 (0.785)	1.818*** (0.638)	2.162** (0.944)	
Month	0.176 (0.676)	0.668 (0.629)	1.274* (0.718)	0.32 (0.569)	1.032** (0.430)	
Nonpara	-0.257** (0.120)	-0.277** (0.119)	-0.271** (0.111)	0.2278*** (0.107)	-0.275*** (0.107)	
Interview	1.577*** (0.506)	0.145 (0.738)	0.970 (0.889)	-0.701 (0.820)	0.148 (0.587)	
Mandatory	0.196 (0.121)	0.203* (0.120)	0.241** (0.113)	0.208* (0.109)	0.220** (0.108)	
Turtle		-0.417 (0.530)	-0.710 (0.509)	-1.013*** (0.326)	-0.867** (0.386)	0.004 (0.469)
Mammal		1.748** (0.862)	0.788 (0.844)	1.498** (0.595)	1.620*** (0.595)	2.574*** (0.470)
Australia			1.034 (0.937)	-1.723* (0.963)		
Philippin			-0.987*** (0.228)	-0.141 (0.332)		
Southeast					-0.270 (0.280)	-0.587** (0.288)
EEPSEA					-2.768 (2.001)	-0.441 (1.720)
LnIncome				0.876*** (0.266)	0.803*** (0.225)	0.505** (0.228)
LnYear				2.145*** (0.611)	3.848** (2.074)	-0.669 (1.445)
Summary statistics						
R ² : within	0.077	0.077	0.190	0.253	0.255	0.168
R ² : between	0.573	0.764	0.881	0.961	0.953	0.727
R ² : overall	0.354	0.500	0.755	0.816	0.813	0.611
Sigma_u	0.801	0.712	0.674	0.386	0.442	0.630
Sigma_e	0.464	0.464	0.437	0.421	0.420	0.435
Rho	0.748	0.701	0.704	0.457	0.525	0.677
N	124	124	124	124	124	124
#Studies	16	16	16	16	16	16

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. STATA 9.2 used

Blank space means variable not included in regression

particular interest for testing in BT how ignoring methodological differences translates into BT errors predicting values for new sites. Note that the models are directly comparable since they all include the same observations.

Individual parameter estimates in the best Model 4 conform fairly well to expectations, where such priors exist. The DC format tends to provide higher estimates than other formats, as expected. Non-parametric estimates are significantly lower than estimates using parametric methods, also as expected. Household payment is significantly higher than WTP from individual payment, though theoretical and empirical expectations here are not clear. Personal interview is not significantly different than other survey modes in the more fully specified Models 2–5. The mandatory payment vehicle tends to give higher WTP (for Models 2–5). It is not immediately clear from our source studies why we find this result. Jianjun et al. (2008) suggest that if CV respondents are assumed to answer truthfully (as if faced with a true economic choice to contribute voluntarily), free riding effects may predict that stated WTP under a voluntary payment mechanism will be lower than when payments are to be made mandatorily (Champ et al. 2002; Wisser 2007).

Valuation of turtle preservation is significantly lower than that for other species (though insignificant in Models 2, 3 and 6), while mammals are valued significantly higher in four out of five models.¹⁷ Higher values for mammals can be explained by their higher degree of “charisma” than for other, lower-profile species. Australian studies provide lower WTP estimates than other countries, when controlling for income level in Models 4. This may be because Australian studies value species we have classified as “non-charismatic”, i.e. the possum (see supplementary appendix).

Studies conducted in the Philippines are likely to give lower values (though only significant in Model 3) than studies conducted in other countries. The income parameter, i.e. the income elasticity of WTP in our double-log formulation, is around 0.5–0.8 and highly significant in the best Models 4 and 5. Income elasticity of WTP in the 0–1 range is commonly found in the CV literature (e.g. Kriström and Riera 1996; Schläpfer 2006). In Model 4 more recent studies yield significantly lower WTP estimates, reflecting perhaps increased prudence in the use of valuation methods over time.

4.2 Level 2: Biodiversity and nature conservation more generally

In Table 7 we present results of five random-effects GLS regression models using the more heterogeneous Level 2 data (nature and biodiversity conservation, i.e. Level 1 data are included as part of the Level 2 data). In this case we include the fuller range of explanatory variables (e.g. covering different valuation methods) using different subsets of the data. We keep the same methodological variables (except we include the dummy for stated preference values¹⁸) for the sake of comparing the robustness of the results with Level 1. Further, we include the

¹⁷ We also tried other groupings or specifications of types of species, such as size, degree of “charisma” across types of species etc., but found that using the biological classification “mammal” worked best in our models. Adding dummies for each species is not feasible due to the limited number of observations for each.

¹⁸ The variable mandatory is excluded here as it is collinear with the valuation method variables.

Table 7 Meta-regression models for Level 2: Biodiversity and nature conservation (standard error of coefficients in parentheses)

Variable ^a	Model 1 GDP inserted for income	Model 2 Income reported	Model 3 Only SP and TCM	Model 4 All variables	Model 5 No method variables
Constant	3.455** (1.513)	4.058*** (1.221)	3.448*** (1.104)	6.554*** (1.800)	5.522*** (1.664)
SP	-0.450 (0.321)	-1.713*** (0.279)	-1.769*** (0.225)	-2.593*** (0.628)	
DC	0.580*** (0.213)	0.011 (0.181)	-0.065 (0.140)	0.760*** (0.221)	
TCM				-2.657*** (0.676)	
Hhdipay	0.335 (0.290)	0.025 (0.260)	0.008 (0.273)	-0.085 (0.332)	
Month	0.606 (0.376)	1.377*** (0.309)	1.448*** (0.310)	1.021** (0.404)	
Nonpara	-0.252 (0.243)	-0.209 (0.174)	-0.220* (0.125)	-0.267 (0.237)	
Interview	0.080 (0.298)	-0.009 (0.249)	0.176 (0.220)	0.153 (0.309)	
Turtle	-0.026 (0.665)	-0.117 (0.490)	-0.275 (0.495)		
Mammal	1.666*** (0.614)	1.885*** (0.494)	1.715*** (0.495)		
Marine	0.888*** (0.308)	0.562** (0.266)	0.554** (0.272)	0.046 (0.447)	0.134 (0.437)
Wetland	-0.991** (0.429)	1.258*** (0.421)	1.218*** (0.414)	-1.967*** (0.538)	-1.718*** (0.528)
Species				-0.942** (0.439)	-0.372 (0.423)
Terrestrial				-1.143*** (0.446)	-0.893* (0.442)
Nonuse	0.057 (0.237)	-0.240 (0.217)	-0.084 (0.179)	0.175 (0.237)	0.093 (0.210)
Oceania	0.755* (0.458)	0.677* (0.405)	0.588 (0.405)	0.994 (0.647)	0.513 (0.630)
East	-0.204 (0.413)	0.180 (0.356)	-0.105 (0.370)	-0.421 (0.638)	-0.646 (0.632)
Southeast	-0.766* (0.412)	-0.323 (0.356)	-0.841*** (0.382)	-0.879 (0.670)	-0.975 (0.665)
South				0.131 (0.751)	0.433 (0.731)
EEPSEA	-0.449** (0.389)	-0.561* (0.310)	0.188 (0.368)	-0.357 (0.403)	-0.266 (0.406)
Journal	-0.318 (0.341)	-0.263 (0.304)	-0.017 (0.318)	-0.096 (0.371)	-0.354 (0.366)
LnIncome	-0.022 (0.133)	0.062 (0.107)	0.103 (0.091)	-0.027 (0.140)	-0.068 (0.136)
LnYear	0.281 (0.236)	0.213 (0.193)	0.180 (0.189)	0.168 (0.262)	0.020 (0.256)

Table 7 continued

Variable ^a	Model 1 GDP inserted for income	Model 2 Income reported	Model 3 Only SP and TCM	Model 4 All variables	Model 5 No method variables
Summary statistics					
R^2 : within	0.124	0.124	0.212	0.179	0.103
R^2 : between	0.172	0.550	0.572	0.155	0.074
R^2 : overall	0.159	0.337	0.459	0.135	0.095
Sigma_u	0.955	0.708	0.764	1.032	1.037
Sigma_e	1.083	0.809	0.582	1.066	1.108
Rho	0.437	0.434	0.632	0.484	0.466
<i>N</i>	550	431	390	550	550
#Studies	95	70	67	95	95

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. STATA 9.2 used

^a Blank space means variable not included in regression

habitat/good characteristics variables that are significant across at least one of our four models. Finally, geographic region dummies were included if significant or if data from these regions dominate our dataset.

Similar to the models in Table 6, we first run a fully specified Model 4 using all variables in Tables 1, 2, 3, 4 and then we exclude in Model 5 method variables. Model 1 investigates the full dataset of 550 observations, inserting GDP as proxy for unreported income information, while Model 2 excludes studies that did not report income information. These two models illustrate the difference between a “conservative” meta-analyst who would not accept the imprecision introduced by inserting proxies for unreported variables and a more “pragmatic” approach. Both practices are found in the MA literature. Model 3 contains the Model 2 observations, excluding values estimated using other methods than CV, CM and TCM (i.e. market price and HPMs¹⁹). Excluding such values altogether is an alternative to using covariates (most common approach in the MA literature), as these methods estimate conceptually different welfare measures.

Given the heterogeneity of the good included in the Level 2 data, our fully specified Model 4 does not do very well in controlling for this heterogeneity using the covariates we have coded and included. The model explains only 13.5% of the variation. This is only slightly increased for Model 1, to a R^2 of 16%, which offers the best combination of covariates for the full dataset. However, it is comparable to the 25–26% obtained in two national level MA studies of an apparently more homogenous good; recreation activity days in the USA (see Rosenberger and Loomis 2000a; Shrestha and Loomis 2003).²⁰ Our R^2 for the full dataset is generally higher than the random-effects MA models of international biodiversity studies in Jacobsen and Hanley (2009).

Excluding the studies from Model 1 for which a crude GDP/capita measure was substituted for missing income information more than doubles the explained variation (Model 2, $R^2 = 0.34$). The coefficient on income turns positive, but is not significant as is the case also for the other models. Hence, the income variable is only significantly positive for the Level 1 data. Insignificant income variable is commonly observed in the MA literature also for less heterogeneous data than our Level 2 data. Enhancing methodological homogeneity in Model 3 increases the explained variation further to 46%, the same level as for example found in the MA of Brander et al. (2006) of international wetland valuation studies.

Despite a higher degree of heterogeneity than for the Level 1 dataset, the data show some degree of regularity, and many of the parameters have the expected signs. Stated preference (SP) methods tend to give lower estimates than revealed preference (RP) methods, as expected (see, e.g. Carson et al. 1996). It is also as expected that monthly payments yield higher estimates than other payment vehicles (Models 2–4) and that non-parametric estimates are lower than parametric ones, like for the Level 1 dataset. Variables such as household’s WTP and in-person interview

¹⁹ The TCM variable is the “hidden” category in Model 3, now that other non-SP methods are excluded. In Models 1–2 the TCM variable is excluded as it is not significant across models.

²⁰ Since R^2 obtained from random-effects models is not directly comparable to standard R^2 OLS, the comparison should be interpreted with caution.

do not significantly influence WTP in any models. Non-parametric estimates are significantly lower than estimates derived by parametric methods only in Model 3. The signs and significance of the mammal coefficients are preserved from the Level 1 models, while the turtle variable is no longer significant in any of the models.

Marine habitats are valued significantly higher than other habitats across Models 1–3, while the wetland coefficient is not robust across models. Value estimates that are primarily non-use related are not significant across models. Studies conducted in Oceania (mostly Australia) tend to yield significantly higher values (significant in Models 1–2). Studies from Southeast Asia give lower values in Models 1 and 3, but are not significant in the other models. Interestingly, studies funded by EEPSEA give lower values than studies funded by other institutions, though not robust across all models. Year is positive but not significant in any models. Removing the methodological variables from the fully specified Model 4 reduces the explanatory power to a low 9.5% in Model 5—an aspect that may invalidate the model for BT purposes.

Increasing the degree of homogeneity of our data in terms of good characteristics and methods, then, generally increases the conformity with theoretical and empirical expectations and explanatory power of the models, as expected. For the more homogenous Level 1 data, observable characteristics of the type of species, region and other variables (income, year) add significantly to the explanatory power of the models. Even with the fairly heterogeneous Level 2 dataset, two models are still able to explain a significant part of the variation of up to 34 and 46%, respectively. This is comparable with other MA studies. For example, the R^2 of 46% of the Level 2 Model 3 is only somewhat lower than what was found in the MA of van Houtven et al. (2007) of water quality valuation studies in the USA. They screened 300 publications related to water quality valuation and found only 11 studies (96 observations) they considered “sufficiently comparable” to be included in the MA. However, for our most heterogeneous Models 1 and 4, the chosen covariates are not able to control for the heterogeneity in a satisfactory way, judged from the level of explained variation.

5 Discussion and conclusions

Pushing the boundaries of meta-analysis (MA) in environmental economics, we have taken stock of studies estimating WTP for conservation of species, biodiversity and nature more generally in Asia and Oceania. Our literature review shows that nature conservation is highly valued, probably more so in many cases than the opportunity costs of increasing conservation efforts in the region, though such a comparison is beyond the scope of our study. The valuation literature in the region covers a wide range of methods and goods, displaying increasing degree of methodological sophistication over time.

Dividing our dataset into two levels of heterogeneity in terms of good characteristics and valuation methods, we show that the degree of regularity and conformity with theory and empirical expectations as well as the explanatory power of our MA models is higher for the more homogenous dataset of species values, as

expected. In fact, though the species are different, the values to preserve them generally follow predictable patterns. For example, we find that mammals are generally valued higher than other species, likely due to the “charismatic” nature of this family. Further, WTP increases significantly with income (elasticity is around 0.5–0.8). The analysis of the species data show that around half of the variation in the best model is due to non-study specific observable characteristics of the good and population surveyed, boding well for use of such data in benefit transfer (BT) applications. However, importantly, increasing the scope of the MA, i.e. gradually including more heterogeneous observations, generally preserves some of the results: The explanatory power of some of our models is in the range of other MA studies of goods typically assumed to be more homogenous (such as national water quality, recreation days, etc.).

Using MA for BT involves transferring one or more estimated meta-regression equations to an unstudied policy site and inserting values from this site to the geographic, socio-economic, good characteristics variables and relevant year and predict or forecast annual WTP per household.²¹ In a preliminary BT investigation subjecting both our dataset levels to a simple check of so-called transfer errors (TE),²² using the MA models to predict observations one-by-one when excluded from the datasets, show for the best models median (mean) TE of 24 (46)% for the species data and 46 (89)% for the more heterogeneous nature and biodiversity data (results in Tuan and Lindhjem 2009). In this test, the estimates left out for the MA model to predict work as benchmarks for the “true value” at a policy site while the model prediction is the transferred value. The transfer errors we found are in the low range compared with other MA studies (see e.g. Lindhjem and Navrud 2008). These preliminary results suggest that such levels of forecasting errors may approach acceptable levels for policy use. It is, however, also clear from our results in this paper that for example including values estimated using a more heterogeneous set of methods for the Level 2 data, even a fairly broad range of covariates is unable to explain and control for the variation in a satisfactory way, potentially translating into large mean transfer errors.

A more careful testing of explanatory variables and MA models for example including interaction effects, isolating heterogeneity along more dimensions and better specifying or capturing the scope dimension of the valued good, all topics for ongoing research, may be required to better understand if heterogeneous good and method characteristics can be controlled for using classical meta-regression analysis. However, even the most extensive list of explanatory variables in the MA-BT literature we have seen to date in Stapler and Johnston (2009) is still unable to bring mean TE below 150% for recreational angling values in North America.

Hence, we are still grappling with the question of how to strike the right balance between screening out studies from the analysis and coding them with the aim of

²¹ The values of methodological variables would typically be set at some best practice level, at the average sample value (Stapler and Johnston 2009) or drawn from the MA sample distribution (Johnston et al. 2006), since there is no such information for an unstudied policy site.

²² $TE = \frac{|WTP_T - WTP_B|}{WTP_B}$, where T = transferred (predicted) value from study site(s), B = estimated (observed) true value (“benchmark”) at policy site.

later controlling for increased heterogeneity in regression models. How homogenous is homogenous enough? Fundamentally, there is still much we do not know about people's preferences and how to represent and interpret them in MA models. Increasing clarity and transparency of effect size definitions, data collection and screening protocols offering others the chance to replicate results, is one important way forward for MA (e.g. as pointed out by Nelson and Kennedy 2009 and USE PA 2006). Using sensitivity analysis to investigate the effects of important analyst choices related to the scope and heterogeneity of the MA dataset is another, as exemplified in this study. This paper is, to our knowledge, one of the first attempts to systematically investigate the issue of heterogeneity in MA of environmental valuation. More research for other goods and geographical areas is needed to inform the development of a more consistent and generally applicable MA methodology, especially as MA is gradually being applied for BT to inform policy. Use of MA in economics is growing and the aim should be to move more of the methodological choices out of the black box.

Acknowledgments We would like to thank Vic Adamowicz, Ståle Navrud and Randall Rosenberger for constructive comments. Funding from the Environment and Economy Programme for Southeast Asia (EEPSEA) is greatly appreciated.

References

- Bandara R, Tisdell C (2005) Changing Abundance of Elephants and Willingness to Pay for their Conservation. *J Environ Manag* 76:47–59
- Barrio M, Loureiro M (2010) A meta-analysis of contingent valuation forest studies. *Ecol Econ* 69:1023–1030
- Bateman JJ, Jones AP (2003) Contrasting conventional with multi-level modeling approaches to meta-analysis: expectation consistency in UK woodland recreation values. *Land Econ* 79:235–258
- Bateman JJ, Carson RT, Day B, Hanemann WM, Hanley N, Hett T, Jones-Lee M, Loomes G, Mourato S, Ozdemiroglu E, Pearce DW, Sugden R, Swanson T (2002) *Economic valuation with stated preference techniques: a manual*. Edward Elgar Publishing, Cheltenham
- Bergstrom JC, Taylor LO (2006) Using meta-analysis for benefits transfer: theory and practice. *Ecol Econ* 60:351–360
- Brander LM, Florax RJGM, Vermaat JE (2006) The empirics of wetland valuation: a comprehensive summary and a meta-analysis of the literature. *Environ Resour Econ* 33:223–250
- Brander LM, van Beukering P, Cesar H (2007) The recreational value of coral reefs: a meta-analysis. *Ecol Econ* 63:209–218
- Carson RT, Flores NE, Martin KM, Wright JL (1996) Contingent valuation and revealed preference methodologies: comparing the estimates for quasi-public goods. *Land Econ* 72:80–99
- Champ PA, Flores NE, Brown TC, Chivers J (2002) Contingent valuation and incentives. *Land Econ* 78:591–604
- Desvousges WH, Johnson FR, Banzhaf HS (1998) *Environmental policy analysis with limited information: principles and applications of the transfer method*. Edward Elgar, Cheltenham
- Engel S (2002) Benefit function transfer versus meta-analysis as policy-making tools: a comparison. In: Florax RJGM, Nijkamp P, Willis KG (eds) *Comparative environmental economic assessment*. Edward Elgar, Cheltenham, pp 133–153
- Freeman AMI (2003) *The measurement of environmental and resource values*. Resource for the Future Press, Washington
- Fromm O (2000) Ecological structure and functions of biodiversity as elements of its total economic value. *Environ Resour Econ* 16:303–328

- Glass GV, McGaw B, Smith ML (1981) *Meta-analysis in social research*. Sage, Beverly Hills
- Hoehn JP (2006) Methods to address selection effects in the meta regression and transfer of ecosystem values. *Ecol Econ* 60:389–398
- Jacobsen JB, Hanley N (2009) Are there income effects on global willingness to pay for biodiversity conservation. *Environ Resour Econ* 43:137–160
- Jacobsen JB, Boiesen JH, Thorsen BJ, Strange N (2008) What's in a name? The use of quantitative measures versus 'Iconised' species when valuing biodiversity. *Environ Resour Econ* 39:247–263
- Jianjun J, Zhishi W, Xuemin L (2008) Valuing black-faced spoonbill conservation in Macao: A policy and contingent valuation study. *Ecol Econ* 68:328–335
- Johnston RJ, Besedin EY, Iovanna R, Miller CJ, Wardwell RF, Ranson MH (2005) Systematic variation in willingness to pay for aquatic resource improvements and implications for benefit transfer: a meta-analysis. *Can J Agric Econ* 53:221–248
- Johnston RJ, Besedin EY, Ranson MH (2006) Characterizing the effects of valuation methodology in function-based benefits transfer. *Ecol Econ* 60:407–419
- Krström B, Riera P (1996) Is the income elasticity of environmental improvements less than one? *Environ Resour Econ* 7:45–55
- Lindhjem H (2007) 20 years of stated preference valuation of non-timber benefits from Fennoscandian forests: a meta-analysis. *J Forest Econ* 12:251–277
- Lindhjem H, Navrud S (2008) How reliable are meta-analyses for international benefit transfers? *Ecol Econ* 66:425–435
- Lindhjem H, Navrud S (2009) Asking for individual or household willingness to pay for environmental goods: implication for aggregate welfare measures. *Environmental and resource economics* 43:11–29
- Lindhjem H, Navrud S (2011a) Are Internet surveys an alternative to face-to-face interviews in contingent valuation? *Ecol Econ* 70:1628–1637
- Lindhjem H, Navrud S (2011b) Using Internet in Stated Preference Surveys: a review and comparison of survey modes. *Int Rev Environ Resour Econ* (Forthcoming)
- Metrick A, Weitzman ML (1996) Patterns of behavior in endangered species preservation. *Land Econ* 72:1–16
- Millennium Ecosystem Assessment (2005) *Synthesis report*. Island Press, Washington
- Moeltner K, Boyle K, Paterson RW (2007) Meta-analysis and benefit transfer for resource valuation—addressing classical challenges with Bayesian modeling. *J Environ Econ Manag* 53:250–269
- Navrud S, Ready R (eds) (2007) *Environmental value transfer: issues and methods*. Springer
- Nelson JP, Kennedy PE (2009) The use (and Abuse) of meta-analysis in environmental and natural resource economics: an assessment. *Environ Resour Econ* 42:345–377
- Pearce D (2005) *Economists and biodiversity conservation: what can we contribute?*. EAERE, Bremen
- Ready R, Navrud S (2006) International benefit transfer: methods and validity tests. *Ecol Econ* 60:429–434
- Richardson L, Loomis J (2009) The total economic value of threatened, endangered and rare species: an updated meta-analysis. *Ecol Econ* 68:1535–1548
- Rosenberger RS, Johnston RJ (2009) Selection effects in meta-analysis and benefit transfer: avoiding unintended consequences. *Land Econ* 85:410–428
- Rosenberger R, Loomis J (2000a) Panel stratification in meta-analysis of economic studies: an investigation of its effects in the recreation valuation literature. *J Agric Appl Econ* 32:131–149
- Rosenberger RS, Loomis JB (2000b) Using meta-analysis for benefit transfer: in-sample convergent validity tests of an outdoor recreation database. *Water Resour Res* 36:1097–1107
- Schläpfer F (2006) Survey protocol and income effects in the contingent valuation of public goods: a meta-analysis. *Ecol Econ* 57:415–429
- Shrestha RK, Loomis JB (2003) Meta-analytic benefit transfer of outdoor recreation economic values: testing out-of-sample convergent validity. *Environ Resour Econ* 25:79–100
- Smith VK, Pattanayak SK (2002) Is meta-analysis a Noah's Ark for non-market valuation? *Environ Resour Econ* 22:271–296
- Stanley TD, Jarrel SD (2005) Meta-regression analysis: a quantitative method of literature surveys. *J Econ Surveys* 19:299–308
- Stapler RW, Johnston RJ (2009) Meta-analysis, benefit transfer, and methodological covariates: implications for transfer error. *Environ Resour Econ* 42:227–246
- Tuan TH, Lindhjem H (2009) Meta-analysis of nature conservation values in Asia and Oceania: data heterogeneity and benefit transfer issues. In: Lindhjem H (ed) *Methodological issues in meta-*

- analysis, benefit transfer and environmental valuation. PhD Thesis. Department of Economics and Resource Management, Norwegian University of Life Sciences, Norwegian
- USEPA (2006) Report of the EPA work group on VSL meta-analysis, report NCEE-0494. National Center for Environmental Economics, EPA, Washington
- Van Houtven G, Powers J, Pattanayak SK (2007) Valuing water quality improvements using meta-analysis: is the glass half-full or half-empty for national policy analysis? *Resour Energy Econ* 29:206–228
- Wiser R (2007) Using contingent valuation to explore willingness to pay for renewable energy: a comparison of collective and voluntary payment vehicles. *Ecol Econ* 62:419–432
- Zandersen M, Tol RSJ (2009) A meta-analysis of forest recreation values in Europe. *J Forest Econ* 15:109–130