Meta-analysis of nature conservation values in Asia and Oceania:

Data heterogeneity and benefit transfer issues

Tran Huu Tuan[a] and Henrik Lindhjem[bc*]

[a] Hue College of Economics, Hue University, 100 Phung Hung Street, Hue City, Vietnam.

[b] Department of Economics and Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Ås, Norway.

[c] Econ Pöyry, P.O. Box 5, N-0051, Oslo, Norway.

Date of draft: February 2009

* Corresponding author. E-mail: henrik.lindhjem@poyry.com.

Meta-analysis of nature conservation values in Asia and Oceania: Data heterogeneity and benefit transfer issues

Abstract

We conduct a meta-analysis (MA) of around 100 studies valuing nature conservation in Asia and Oceania. Dividing our dataset into two levels of heterogeneity in terms of good characteristics (endangered 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 endangered 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. Subjecting our best MA models to a simple benefit transfer test forecasting values for out-of-sample observations, shows median (mean) transfer errors of 24 (46) percent for endangered species and 46 (89) percent for nature conservation more generally, approaching levels that may be acceptable in benefit transfer for policy use. However, as for example more heterogeneous observations are included, our meta-regression models are unable to control for the variation in a satisfactory way, resulting in high mean transfer errors. Despite some encouraging results, more research is clearly required to answer the question of how homogenous is homogenous enough in meta-analysis and benefit transfer.

Keywords: Asia; benefit transfer; biodiversity; meta-analysis; Oceania; valuation

JEL Classification: Q26; Q51; Q57; H41

Introduction

According to the Millennium Ecosystem Assessment, more than 60 per cent of the world's ecosystems are being degraded or used unsustainably (MEC 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. 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). For a range of environmental goods meta-analysis (MA) techniques have been used to synthesize 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). Responding to Pearce's challenge, we use MA to review and take stock of the literature to date on environmental valuation of a complex and somewhat heterogeneous good: (changes in) conservation of habitat, biodiversity and endangered species, in a specific geographical region: Asia and Oceania. We attempt to answer the following two research questions; (1) To what extent do welfare estimates for this complex good

[1] Originally quoted in Stanley and Jarrel (2005).

conform with theoretically and empirically derived expectations regarding the good characteristics, valuation methods, study quality, socio-economic and other variables?; (2) How sensitive are the meta-regression results and the value forecasts for unstudied sites 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; Rosenberger and Johnston 2007). An important analyst 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 between the number of observations and the acceptable level of heterogeneity in the data, as pointed out by e.g. Engel (2002) and Nelson and Kennedy (2009) (Question 2a above). Another, related choice is which model to choose for BT, for example which covariates to include and how to treat insignificant variables (Question 2b)[2]. There are different practices and little is known of the empirical effects of these choices, though Lindhjem and Navrud (2008a) have shown that the precision of meta-analytical BT (MA-BT) depends on the model specifications, sometimes in unexpected ways.

[2] An alternative approach to dealing with classical MA challenges, not pursued here, is to use Bayesian techniques (e.g. Moeltner et al 2007 and Moeltner and Rosenberger 2008).

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 (e.g. wetlands, coral reefs, forests, biodiversity, agricultural land preservation) (see e.g. Brander et al 2006; 2007; Jacobsen and Hanley 2009; Johnston et al 2008; Lindhjem 2007; Lindhjem and Navrud 2008a; Loomis and White 1996; Richardson and Loomis 2009; Stapler and Johnston 2009). With very few exceptions, these studies do not focus specifically on MA methodology or implications for BT, despite the growing focus on meta-analytic BT in the literature. Compared to previous work, we add several new and interesting dimensions; (1) To investigate the effect of MA scope, we divide our dataset into two levels of heterogeneity; endangered species (more similar good and methods used) and nature conservation more generally (more heterogeneity in good and methods used); (2) 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, and finally; (3) We report the level of forecast (or transfer) errors for unstudied sites broken down by type of models, nature conservation habitat, geographic region and valuation method used, based on a jackknife resampling technique used in MA by e.g. Brander et al (2006) and Lindhjem and Navrud (2008a). This study is, to our knowledge, one of the first attempts to systematically investigate the issue of heterogeneity in MA and BT in environmental valuation

Conceptual framework and data

Conceptual framework

We start by defining "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). From the standard indirect utility function, the bid function for a representative individual j for this change can be given by (Bergstrom and Taylor 2006)[3]:

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

Where P = a price index of market goods (assumed constant), M = (individual or household) income (assumed constant), Q^T-Q^R and $QUAL^T$-$QUAL^R$ are the changes in quantity and quality from a reference situation ("status quo") (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 estimates of mean WTP from different

[3] 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.

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, 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[4]. 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.

Description of the meta-data from nature conservation studies

[4] The ecosystem services and functions and total economic value from nature and biodiversity conservation are discussed in depth elsewhere, and for sake of brevity not elaborated in detail here (see e.g. Fromm 2000).

Given this conceptual framework, we conducted a broad search for studies (published papers, reports, book chapters etc[5]) internationally available in English valuing nature conservation in the region drawn from various databases. 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 since the 1990s. 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 procedure 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 are given in Tables 1-3 below[6]. 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 (10). 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).

---

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

[6] We do not claim to have collected an exhaustive database of all studies in Asia and Oceania, but the extent of our search makes us confident that we cover the majority of such studies in the region. Further, it is unlikely that our search has been biased in any way (except for the focus on studies in English). Finally, to answer our research questions, completeness is also not strictly necessary.

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

Region	Mean WTP (St.dev.)	No. of obs.	No. of studies
Southeast Asia (SEA)	83 (212)	244	32
Oceania (O)	254 (914)	116	23
East Asia (EA)	76 (108)	99	23
South Asia (SA)	206 (286)	70	11
Southwest Asia (SWA)	66 (78)	21	6
Total		**550**	**95**

Note: O= Australia, Micronesia, Papua New Guinea, Vanuatu; SEA= Cambodia, Indonesia, Laos, Malaysia, the Philippines, Singapore, Thailand, Vietnam; EA= China, Japan, Korea, Taiwan; SA= India, Sri Lanka; SWA= Iran, Israel, Pakistan.

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
Endangered 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***

Note: * 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***

Note: * In some studies, there are more than one method used, and thus the number of studies is higher than reported in Table 1.

The most frequently valued habitat is terrestrial habitats (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, watercourses) for simplicity termed “marine habitats”, follow second. 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 and endangered species or groups of species are categorised as “endangered species”. Marine habitats provide the lowest value (US$ 80) compared to other types of habitats, while terrestrial habitats (US$ 116), endangered species (US$ 105), and other habitats (US$ 121) have values that are around 40-50 percent higher. The by far most frequently used method is contingent valuation (CV), with 77 studies, while the travel cost method (TCM) comes second with 14 studies (Table 3). A small number of studies (5) use other methods, such as the hedonic pricing method or calculating the value of wetlands and forests using the market price approach. These methods frequently calculate a different welfare measure than CV, CM and TCM studies, and also yield twice as high estimates as the other methods. Details of the individual studies (including references) are given in Appendix A.

Coding of the 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 a large universe of potential 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 give the variable names and definitions. 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[7], 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, we standardized the values to a common metric, 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 per visit WTP 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) exchange rates were used to correct for differences in price levels between countries, which 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 2007; 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), and whether the CV data were analysed using non-parametric statistical methods. 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

---

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

(a main reason is so-called yea-saying), while non-parametric methods typically give a lower bound on WTP[8]. There is no clear prior for use of interviews vs. other modes, though type of survey mode is known to influence results (Lindhjem and Navrud 2008b).

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 vs. 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), 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 are likely to give lower value estimates. There are no strong priors regarding other habitat types or regional/country dummies, though it is expected that these dimensions may influence WTP[9]. 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 percent 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

---

[8] 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. USEPA 2007).

[9] 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.

other sources instead (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 (see e.g. Johnston 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, 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 Stanley 2007). A way to limit the potential impact of publication bias is also to include unpublished studies. To capture trends in WTP values over time 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 by the same institution, Environment and Economy Program for Southeast Asia, 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.

Table 4 Definition of meta-analysis variables and descriptive statistics

Variables	Description	Mean (SD)*
Dependent variable:		
WTP 2006	WTP in 2006 prices (US$)	133 (461)
Methodological variables:		
SP	Binary: 1 if stated preference, 0 if otherwise	.84 (.35)
DC	Binary: 1 if SP using dichotomous choice, 0 if otherwise	.51 (.50)
TCM	Binary: 1 if travel cost method, 0 if otherwise	.07 (.25)
Hholdpay	Binary: 1 if household's WTP, 0 if individual	.67 (.46)
Month	Binary: 1 if payment is a monthly payment, 0 if otherwise	.35 (.47)
Nonpara	Binary: 1 if estimate is non-parametric (Turnbull), 0 otherwise	.07 (.25)
Interview	Binary: 1 if it is an in-person interview, 0 otherwise	.60 (.48)
Good characteristics variables:		
Mammal	Binary: 1 if it is a mammal, 0 otherwise	.04 (.20)
Turtle	Binary: 1 for sea turtle, 0 otherwise	.06 (.24)
Species	Binary: 1 for endangered species, 0 if other habitats/services	.23 (.42)
Terrestrial	Binary: 1 for terrestrial habitats, 0 if other habitats/services	.32 (.47)
Marine	Binary: 1 if marine habitat (beach, sea, watercourse, lake, river), 0 other habitats/services	.29 (.45)
Wetland	Binary: 1 for wetlands, 0 if other habitats/services	.07 (.26)
Nonuse	Binary: 1 for primarily non-use, 0 otherwise	.77 (.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	.19 (.39)
Philippin	Binary: 1 if a study in the Philippines, 0 otherwise	.22 (.42)
Oceania	Binary: 1 if a study in Oceania, 0 other region	.21 (.40)
East	Binary: 1 if a study in East Asia, 0 other region	.18 (.38)
Southeast	Binary: 1 if a study in Southeast Asia, 0 otherwise	.44 (.48)
Southwest	Binary: 1 if a study in Southwest Asia, 0 otherwise	.04 (.19)
South	Binary: 1 if a study in South Asia, 0 otherwise	.13 (.33)
Other variables:		
EEPSEA**	1 if the study is funded by EEPSEA, 0 otherwise	.39 (.48)
Journal	1 if it is a published paper, 0 otherwise	.47 (.49)
Year	Continuous: from 0 (2006) to 26 (1979)	6.36 (4.07)

Notes: *The Mean (SD) 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.
**EEPSEA = Environment and Economy Program for Southeast Asia.

Meta-regression model and results

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: Endangered species; and Level 2: Biodiversity and nature conservation more generally. The endangered 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 programs for endangered species (e.g. WTP to conserve a viable population of sea turtles)[10].

10 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 (Phillipines), 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). The 16 studies 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 to 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, it can be argued, than many other complex environmental goods studied in MA. Further, as almost all non-textbook goods in general (and environmental goods in particular) are heterogeneous to some degree, it is unclear from theory where to draw the line in practice. All in all the Level 2 dataset contains between 67 to 95 studies and 390 to 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 Appendix A. We will conduct several 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 study[11] (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 equation (1) above, can be written as:

---

[10] 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.

[11] 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.

(2) $WTP_{ij} = \alpha + \sum_{i=1}^{n} \beta_i x_{ij} + \mu_{ij} + \varepsilon_i$

where WTP is the i'th observation from the j'th strata (here study), α is a constant, x_{ij} is a vector of explanatory variables (as defined in Table 4), with a panel effect μ_{ij} and an error ε_i ~N(0, σ_ε^2). We also assume that μ_{ij}, ε_i and x_{ij} are uncorrelated within and across studies. A Breusch-Pagan Lagrange multiplier (LM) 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 LM test is rejected, the random effects model assuming heterogenous 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[12], the Breusch-Pagan LM test showed that a model with equal effects was rejected, confirming the appropriateness of a panel estimation model (χ^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

[12] 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

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.

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

	b Fixed effects model	**B Random effects model**	**b-B**	**S.E.**
Income variable	.0305127	-.0494427	.0799554	.2193994
p> χ^2: 0.7155				

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

Results and discussion

First, we provide results in this section of different meta-regression models for Levels 1 and 2 of our data. Then, in the next section, we use the estimated models to investigate errors in predicted values for unstudied sites (i.e. a BT simulation). Results of six random effects GLS regression models for the Level 1 data (species) 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 is able to deal with the heterogeneity of the data, for the most 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 methods 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 variable), 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 40 percent of the variation ($R^2 = 0.398$), while adding characteristics of the species explain another 14 percent of the variation ($R^2 = 0.536$). Adding country specifics and income and year in Models 3 and 4 help explain another 22-27 percentage 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 to 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 percent. This model is particularly interesting for testing in the next section 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 confirm well with expectations, where such priors exist. The DC format tends to provide higher estimates than other formats, as expected, as do monthly payments compared to other periods of payment. Non-parametric estimates are significantly lower than estimates using parametric methods, also as expected. Household payment is significantly higher than WTP from individual payment (Models 4 and 5), 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.

Table 6 Meta-regression models for Level I: Endangered species studies

Variable$^{\#}$	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
	Method variables	**+ Species types**	**+ Country variables**	**+ Income and year**	**All variables**	**No method variables**
Constant	1.298*	2.413***	1.493	-9.365***	(dropped)	-.455
	(.095)	(.002)	(.156)	(.001)		(3.861)
DC	1.517*	.695	1.187	1.555***	1.294**	
	(.064)	(.374)	(.102)	(.002)	(.566)	
Hholdpay	.855	.038	.563	1.722***	2.222**	
	(.295)	(.961)	(.438)	(.003)	(.922)	
Month	.168	.657	1.116*	.140	1.015**	
	(.810)	(.274)	(.092)	(.788)	(.419)	
Nonpara.	-.259**	-.278**	-.273**	-.281***	-.277**	
	(.032)	(.022)	(.016)	(.010)	(.109)	
Interview	1.525***	.113	.729	-.972	.064	
	(.004)	(.873)	(.375)	(.192)	(.571)	
Turtle		-.363	-.675	-.954***	-.830**	.004
		(.470)	(.151)	(.001)	(.375)	(.469)
Mammal		1.740**	.856	1.569***	1.664***	2.574***
		(.035)	(.277)	(.004)	(.578)	(.470)
Australia			.698	-2.048**		
			(.415)	(.019)		
Philippin.			-.982***	-.126		
			(.000)	(.699)		
Southeast					-.248	-.587**
					(.281)	(.288)
EEPSEA					-11.421***	-.441
					(3.908)	(1.720)
Journal					-14.870***	(dropped)
					(5.394)	
LnIncome				.854***	.798***	.505**
				(001)	(.225)	(.228)
LnYear				2.189***	4.424**	-.669
				(.000)	(2.003)	(1.445)
Summary statistics:						
R^2: within	.044	.044	.157	.231	.234	.168
R^2: betw.	.604	.785	.879	.961	.947	.727
R^2: overa.	**.398**	**.536**	**.757**	**.810**	**.804**	**.611**
Sigma_u	.841	.674	.614	.330	.420	.630
Sigma_e	.470	.470	.444	.425	.424	.435
Rho	.761	.672	.656	.376	.459	.677
N	124	124	124	124	124	124
# studies	16	16	16	16	16	16

Note: *p < 0.10, **p < 0.05, ***p < 0.01. STATA 9.2 used. # Blank space means variable not included in regression.

Valuation of turtle preservation is significantly lower than for other species (though insignificant in Models 2 and 3), while mammals are valued significantly higher[13]. 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 Appendix A). 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 0.8 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). In Model 4 more recent studies yield significantly lower WTP estimates, reflecting perhaps increased prudence in the use of valuation methods over time.

In Table 7 we present results of six random effects GLS regression models instead using the more heterogeneous Level 2 data (nature and biodiversity conservation, i.e. Level 1 are a subset 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 values) for the sake of comparing the robustness of the results with Level 1. Further, we include the 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

---

[13] 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.

Table 6, we first run a fully specified Model 5 using all variables in Tables 1-4 and then we exclude in Model 6 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 illustrates 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 hedonic pricing methods[14]), as these methods typically estimate conceptually different welfare measures. Model 4 contains studies of endangered species only (the same observations as in Model 4 from Level 1), for sake of comparison. Given the heterogeneity of the good included in the Level 2 data, our fully specified Model 5 does not do very well in controlling for this using the covariates we have been able to code and include. The model explains only 13.5 percent of the variation. This is only slightly increased for Model 1, to a R^2 of 16 percent, which offers the best combination of covariates for the full dataset. However, it is comparable to the 25-26 percent obtained in two national level MA studies of an apparently more homogenous good; recreation activity days in the USA (see Rosenberger and Loomis 2000a and Shresta and Loomis 2003)[15]. 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).

[14] 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.

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

Table 7 Meta-regression models for Level 2: Biodiversity and nature conservation

Variable[#]	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	GDP inserted for income	**Income reported**	**Only SP and TCM**	**Level 1 data**	**All variables**	**No method variables**
Constant	3.455**	4.058***	3.448***	(dropped)	6.554***	5.522***
	(.022)	(.001)	(.002)		(1.800)	(1.664)
SP	-.450	-1.713***	-1.769***	(dropped)	-2.593***	
	(.149)	(.000)	(.000)		(.628)	
DC	.580***	.0114	-.065	-1.856	.760***	
	(.007)	(.950)	(.642)	(.163)	(.221)	
TCM					-2.657 ***	
					(.676)	
Hhldpay	.335	.025	.008	-2.270**	-.085	
	(.248)	(.923)	(.976)	(.032)	(.332)	
Month	.606	1.377***	1.448***	2.893***	1.021**	
	(.108)	(.000)	(.000)	(.000)	(.404)	
Nonpara	-.252	-.209	-.220*	-.307***	-.267	
	(.300)	(.229)	(.078)	(.006)	(.237)	
Interview	.080	-.009	.176	1.749**	.153	
	(.778)	(.970)	(.442)	(.049)	(.309)	
Turtle	-.026	-.117	-.275	-.912** (.014)		
	(.968)	(.811)	(.579)			
Mammal	1.666***	1.885***	1.715***	1.710***		
	(.007)	(.000)	(.001)	(.002)		
Marine	.888***	.562**	.554**	(dropped)	.046	.134
	(.004)	(.035)	(.042)		(.447)	(.437)
Wetland	-.991**	1.258	1.218***	(dropped)	-1.967***	-1.718***
	(.021)	(.003)***	(.003)		(.538)	(.528)
Species					-.942**	-.372
					(.439)	(.423)
Terrestrial					-1.143***	-.893
					(.446)	(.442)
Nonuse	.057	-.240	-.084	(dropped)	.175	.093
	(.809)	(.269)	(.639)		(.237)	(.210)
Oceania	.755*	.677 *	.588	(dropped)	.994	.513
	(.099)	(.095)	(.146)		(.647)	(.630)
East	-.204	.180	-.105	-3.825	-.421	-.646
	(.622)	(.612)	(.776)	(.108)	.638	(.632)
Southeast	-.766*	-.323	-.841**	-3.997*	-.879	-.975
	(.063)	(.364)	(.028)	(.080)	(.670)	(.665)
South					.131	.433
					(.751)	(.731)
EEPSEA	-.449**	-.561*	.188	(dropped)	-.357	-.266
	(.024)	(.070)	(.609)		(.403)	(.406)
Journal	-.318	-.263	-.017	-5.309	-.096	-.354
	(.351)	(.387)	(.956)	(.373)	(.371)	(.366)
LnIncome	-.022	.062	.103	.867***	-.027	-.068
	(.863)	(.558)	(.260)	(.000)	(.140)	(.136)
LnYear	.281	.213	.180	.791	.168	.020
	(.234)	(.270)	(.342)	(.818)	(.262)	(.256)
Summary statistics:						
R^2 within	.124	.124	.212	.227	.179	.103
R^2:betwen	.172	.550	.572	.953	.155	.074
R^2:overall	**.159**	**.337**	**.459**	**.790**	**.135**	**.095**
Sigma_u	.955	.708	.764	.396	1.032	1.037
Sigma_e	1.083	.809	.582	.440	1.066	1.108
Rho	.437	.434	.632	.447	.484	.466
N	550	431	390	124	550	550
# studies	95	70	67	15¤	95	95

Note: *p < 0.10, **p < 0.05, ***p < 0.01. STATA 9.2 used. # Blank space means variable not included in regression. ¤ Due to different st.dev. in Level 2 data, screening procedures reduce studies from 16 to 15 for Level 1 data, though number of observations turns out by coincidence to be the same.

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. Enhancing methodological homogeneity in Model 3 increases the explained variation further to 46 percent, the same level as for example found in the MA of Brander et al (2006) of international wetland valuation studies. Finally, in Model 4, using the Level 1 dataset, with the more complete range of explanatory variables does not change R^2 much compared to Model 4 in Table 6. 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. It is also as expected that monthly payments yield higher estimates than other payment vehicles and that non-parametric estimates are lower than parametric ones, like for the Level 1 dataset. The other methodological parameter estimates (i.e. household WTP, personal interview) are not robust across models and there are no strong priors for their signs. The signs and significance of the turtle and mammal parameters are preserved from the Level 1 models.

Marine habitats are valued significantly higher than other habitats across Models 1-3, while wetlands tend to be valued higher (though not robust across models). Primarily non-use value estimates are not significant across models. Studies conducted in Oceania (mostly Australia) tend to yield significantly higher values (in Models 1-2). Studies from Southeast Asia give lower values, compared to other regions. Interestingly, studies funded by EEPSEA give lower values than studies funded by other institutions, though not robust across all models. The income parameter is positive for studies that have reported income information from their samples, but only significantly so in Model 4 for the endangered species data. Year is positive but not significant in any models. Removing the methodological variables from the fully specified Model 5, reduces the explanatory power to a dismal 9.5 percent in Model 6 – an aspect that may invalidate the model for BT purposes (see next section). In addition to the models in Table 6 and 7, we also ran Models 1-4 in reduced form, in which variables not significant at the 20 percent level are removed. This form is often used in MA-BT (see Rosenberger and Loomis 2000a and Lindhjem and Navrud 2008a). Detailed regression results are given in Appendix B. Finally, we also removed the method variables for Model 3 (Level 2), to test if that model may still be useable for BT, if choosing to ignore methodological differences (see regression results in Appendix C). The results are discussed in relation to the BT simulations in the next section.

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 percent, respectively, comparable with other MA studies. For example, the R^2 of 46 percent of the Level 2 Model 3 is only about 10-15 percentage points 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 5, the chosen covariates are not able to control for the heterogeneity in a satisfactory way, judged from the level of explained variation. Given the degree of confirmed validity of our data and explanatory power, the next, and directly policy relevant question, is how the MA models will perform forecasting values for unstudied sites, i.e. used for BT. This is the question we turn to in the next section.

A check of the transferability of nature conservation values

MA-BT involves transferring one or more estimated meta-regression equations (2) to an unstudied policy site, and insert values from this site for the geographic, socio-economic, good characteristics variables and relevant year, and predict or forecast annual WTP per household. 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. To the extent observable characteristics of the habitats/good valued and the population explain a significant portion of the variation in WTP, and not only the methodological differences between studies, it gives us confidence that MA-BT may be a credible alternative to a new valuation study as input for example in cost-benefit analysis. The performance of MA-BT can only be accurately assessed if we knew the "true value", or an estimate of this, for a range of sites, and then used the MA models to predict the value at those sites, and calculate so-called transfer errors (TE)[16]. Lindhjem and Navrud (2008a) and a few other studies, use different "benchmark" values from within their sample or from new studies to "simulate" the true value to assess TE performance. We will not conduct a full such investigation, but only carry out a first check on how our MA models forecast nature conservation values for

[16] $\mathrm{TE} = \frac{|\mathrm{WTP}_T - \mathrm{WTP}_B|}{\mathrm{WTP}_B}$, where T = Transferred (predicted) value from study site(s), B = Estimated (observed) true value ("benchmark") at policy site.

our two datasets. We use a jackknife data splitting technique, used e.g. by Brander et al (2006), where we estimate n-1 separate meta-regression equations to predict (or forecast) the value of the omitted observation in each case (i.e. "the policy site"). We then calculate the percentage difference between observed and predicted values, the TE in our simple exercise, and the overall median and mean TE for all observations[17].

This measure gives a first indication of how far off our MA models would be in a real BT exercise. We start by reporting the results for the six models using the Level 1 and Level 2 data (Table 8 and 9 below, respectively)[18].

Table 8 Median and mean transfer error (percent) for models Level 1: Endangered species

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Method variables	**+ Species types**	**+ Country variables**	**+ Income and year**	**All variables**	**No method variables**
Median	61	59	33	24	23	36
Mean	108	85	58	46	45	67
N	124	124	124	123	123	123

Table 9 Median and mean transfer error (percent) for models Level 2: Biodiversity and nature conservation

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	GDP inserted for income	**Income reported**	**Only SP and TCM**	**Level 1 data**	**All variables**	**No method variables**
Median	71	52	46	22	77	64
Mean	7344	377	89	45	5277	8363
N	547	428	387	121	547	547

First, using a fully specified model such as Models 4 and 5 for Level 1, yields fairly high precision at around median TE of 23-24 percent (mean 45-46 percent) (Table 8).

[17] The mean prediction error is often termed Mean Absolute Percentage Error (MAPE).

[18] To account for econometric error in transforming ln(WTP) to WTP using antilog, we add standard deviation ($s^2/2$), which estimate varies when the sample changes, prior to transformation of ln(WTP) (see e.g. Johnston et al 2006) Some of the observations were dropped by STATA performing the TE estimations in Tables 8-9 as compared to Tables 6-7.

Mean TE is low compared to other studies performing this check, e.g. Lindhjem and Navrud (2008a) (62-266 percent), Brander et al (2006; 2007) (74-186 percent), Stapler and Johnston (2009) (152 percent), Richardson and Loomis (2009) (34-45 percent for a within-sample test), indicating a level of precision that could be acceptable for policy use. A general level of 20-40 percent has been suggested (Kristofersson and Navrud 2007). Precision increases with the more fully specified models, as expected. Interestingly, introducing species dummies to Model 1 reduces median TE by only 2 percentage points in Model 2 compared to an almost halving of TE from introducing country variables in Model 3. Model 6, which uses only observable values at the "policy site" (i.e. no methodological controls included), still manages to predict values with a median precision of 36 percent. Stapler and Johnston (2009) find that using the hypothetical "ideal" values for the methodological variables instead of the means in the sample in a fully specified MA-BT model, gives a TE gain overall of only 26 percentage points (from 151.9 to 125.6 percent). In an earlier study Johnston et al (2006) find that the choice of values for methodological variables may have a large impact on forecasts. Although ignoring methodological differences between studies altogether may not generally be a sensible approach in MA-BT, our case illustrates at least that loss in precision is relatively low as long as the good and methods used are relatively similar. We also ran the BT tests using reduced versions of Models 1-4 (Level 1 & 2), excluding variables not significant at the 20 percent level, a common procedure in MA-BT studies. We found very similar TE levels and no clear relationship between TE and reduced vs. full versions of the models (see results in Tables B1-B4 in Appendix B)[19].

[19] We also ran the same TE simulations using a rule-of-thumb of p>0.1 instead of p>0.2 for the reduced models,

For the Level 2 data Models 1-4 median TE is, somewhat surprisingly, comparable to the Model 1-4 Level 1 results, despite lower explanatory power (Table 9). However, the Level 2 data produce more high TE values (i.e. the mean is much higher than the median). Keeping to the median TE, Models 5 and 6 for Level 2 do a poorer job at controlling heterogeneity as judged by the TEs of 77 and 64 percent, respectively, compared to Models 5 and 6 for Level 1. However, although the mean TEs are in the thousands, it is somewhat surprising that medians are not influenced more by the low explanatory power of Models 5 and 6 for Level 2. We also ran a version of Model 6 for Level 2 based on the more methodologically homogenous dataset used for Model 3. This produced median TE of 49 percent (mean 108 percent), down from 64 percent in Table 6, caused by only excluding hedonic and market price method observations (having no methodological covariates)[20] (see Model 6a in Appendix C for results). Reducing methodological heterogeneity for the Level 2 data from Model 2 to 3 reduces median TE from around 52 to 46 percent, while mean TE comes down from an unacceptably high level of 377 percent to a more reasonable 89 percent. For both Level 1 and 2 models there is generally an inverse relationship between the level of explained variation and TE, as expected. Hence, increasing degree of homogeneity of the data in terms of good characteristics (biodiversity and nature conservation in general to endangered species) increases the precision, as does the enhanced homogeneity of valuation methods used within Level 2. However, in median terms, the gain in precision is perhaps not as highly related to explanatory power or homogeneity, as expected. Even with a heterogeneous dataset, median TE may approach acceptable levels for policy use.

---

detecting no clear(er) relationship with TE.

[20] We acknowledge that reducing methodological heterogeneity may also implicitly reduce good heterogeneity, as some types of nature conservation values may be more likely to be estimated using particular methods.

But for mean TE the results are clear. The plot of observed WTP values (estimates sorted in ascending order) vs. predicted (zigzag line,) for Model 4 (Level 1 data) is illustrated in Figure 1. The forecasts follow the observed values well except at the extremities of the data, a characteristic of forecasting models. For comparison, Model 1 (the whole dataset, 550 observations) for Level 2 is plotted in Figure 2. This plot shows a lower level of precision than for Level 1 in Figure 1 (note that the scale is different).

Figure 1 Plot of predicted vs. observed lnWTP for Model 4 (Level 1)

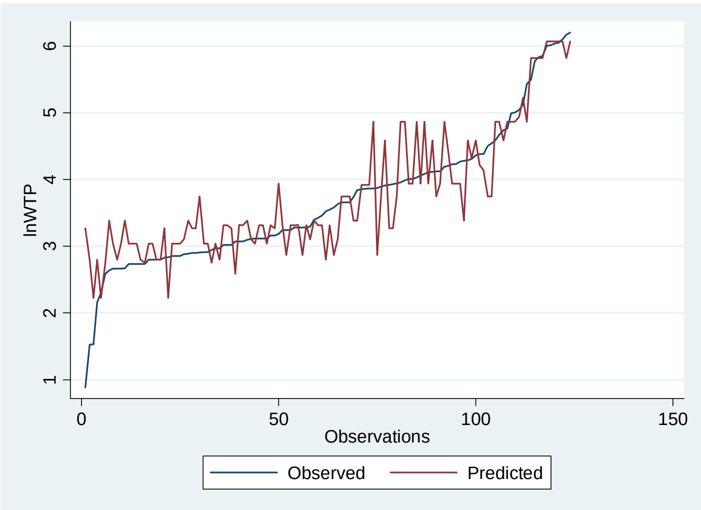
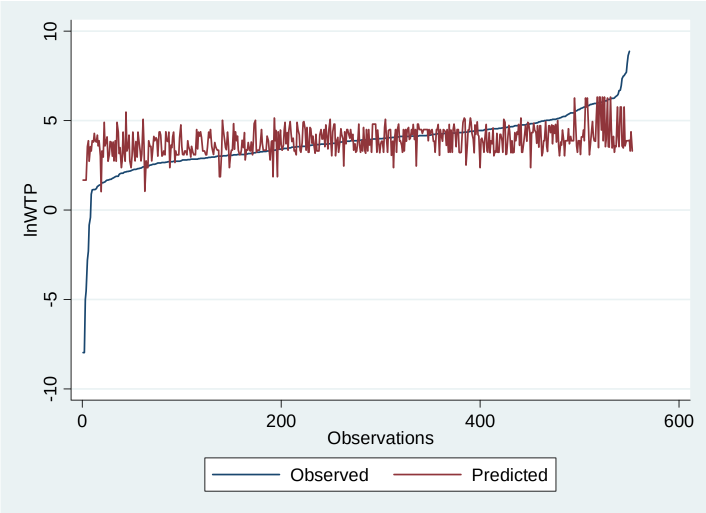


Figure 2   Plot of predicted vs observed lnWTP for Model 1 (Level 2)



We also break down estimated median and mean TE from Tables 8 and 9 for Models 1-4 only, for different subsets of the Level 1 and 2 data, i.e. by different characteristics of the good (Levels 1 and 2), valuation methods and geographical region (Level 2). First, TE for the four Level 1 models predicting values for preservation of turtles, mammals and other species, are given in Table 10. Precision increases from Model 1 through to 4 for all species types. WTP for mammal preservation is predicted with a median (mean) precision of 16 (17), percent, while for other species median TE doubles.

Table 10   Median (mean) transfer error in percent for different species, Models 1-4, Level 1 data

	Turtle	**Mammal**	**Other species**
Model 1	45 (169)	50 (120)	67 (75)
Model 2	52 (114)	43 (86)	65 (71)
Model 3	32 (97)	36 (61)	33 (39)
Model 4	24 (69)	16 (17)	32 (43)
# of obs.	**34**	**19/20**	**70**

In Table 11 we split the estimated TE for species and different types of habitats for the Level 2 data. The precision is generally higher for the endangered species (median TE of 36 percent for Model 3 and 22 percent for Model 4). Model 3 predicts WTP for terrestrial and marine habitats with the same median error of 40-46 percent (means at around 100 percent), while wetlands and other habitats have higher median errors. The similarity (homogeneity) in the change of environmental quality (ensure preservation) among the endangered species studies may be the reason that their values are predicted at a higher level of precision. Santos (1998) argues that the prediction errors he obtains in a MA of CV studies of landscape conservation are higher than those estimated in Loomis' (1992) study of rivers within the same US state, due to landscapes being a more heterogeneous good.

Table 11 Median (mean) transfer error in percent for different habitat types, Models 1-4, Level 2 data

	Terrestrial habitats	**Marine habitats**	**Endangered species**	**Wetlands**	**Other habitats**
Model 1	86 (545)	63 (605)	47 (85)	71 (92838)	77 (184)
Model 2	62 (1134)	44 (105)	36 (60)	77 (116)	78 (79)
Model 3	46 (104)	40 (106)	36 (57)	71 (119)	81 (75)
Model 4	-	-	22 (45)	-	-
# obs.	81- 173	129-162	121-129	31-41	17-37

Enhancing methodological homogeneity from Model 2 to 3 (i.e. removing estimates using market price or hedonic methods) reduces TE especially for terrestrial habitats. This is an indication that other valuation methods introduce substantial noise for terrestrial habitat valuation in the MA. In Table 12 we break down TE by valuation methods used. WTP estimates derived by CV has a median (mean) TE of 41 (71) percent in the most homogenous Model 3. Estimates derived by TCM or other valuation methods generally have higher TE than stated preference methods.

Table 12 Median (mean) transfer error in percent for different valuation methods, Models 1-4, Level 2 data

	CV	CM	TCM	Others
Model 1	64 (157)	70 (159)	73 (104785)	93 (1776)
Model 2	41 (81)	78 (167)	101 (141)	84 (3882)
Model 3	41 (71)	26 (149)	105 (145)	-
Model 4	22 (45)	-	-	-
# of obs.	121-423	50	17-37	37

Finally, breaking the TE estimates down by region in Table 13 shows that using the model to predict values in Southeast Asia produces the lowest TE, which is partly due to the larger number of estimates from this region. Except for some very high TE estimates pulling up the mean, median TE for transfers to all regions is below 80 percent and approaching acceptable levels for policy use.

Table 13 Median (mean) transfer error in percent for different regions, Models 1-4, Level 2 data

	Southeast Asia	Oceania	East Asia	South Asia	Southwest Asia
Model 1	59 (16000)	80 (209)	66 (163)	80 (1605)	45 (76)
Model 2	42 (102)	61 (81)	68 (304)	67 (3184)	44 (43)
Model 3	37 (90)	59 (76)	68 (110)	21 (27)	36 (36)
Model 4	23 (44)	31 (50)	28 (57)	15 (17)	-
# of obs.	69-244	16-116	26-99	10-41	12-21

Concluding remarks

Pushing the boundaries of meta-analysis (MA) in environmental economics, we have taken stock of studies estimating willingness to pay (WTP) for conservation of endangered 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. 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 endangered 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 0.8). The analysis of the endangered 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 regularity and 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).

Subjecting both our dataset levels to a simple check of benefit transfer error (TE), 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) percent for the endangered species data and 46 (89) percent for the more heterogeneous nature and biodiversity data. This is in the low range compared to other MA studies. Results suggest that such levels of forecasting errors may approach acceptable levels for policy use. However, caution should be exercised in using values for single species for BT, as such estimates may include values of biodiversity or habitats more generally (see e.g. Veisten et al 2004). It is also clear from our results 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, translating into large mean transfer errors. A more careful testing of explanatory variables and MA models than we have done (for example including interaction effects) 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 MA-BT we have seen to date in Stapler and Johnston (2009) is still unable to bring mean TE below 150 percent 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 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 USEPA 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 and BT in 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.

Acknowledgements

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 and Tisdell C (2005) Changing Abundance of Elephants and Willingness to Pay for their Conservation. Journal of Environmental Management 76: 47-59

Bateman I J and Jones A P (2003) Contrasting conventional with multi-level modeling approaches to meta-analysis: Expectation consistency in UK woodland recreation values. Land Economics 79(2): 235-258

Bergstrom J C and Taylor L O (2006) Using meta-analysis for benefits transfer: Theory and practice. Ecological Economics 60: 351-360

Brander L M, Florax R J G M and Verrmaat J E (2006) The Empirics of Wetland Valuation: A Comprehensive Summary and a Meta-Analysis of the Literature. Environmental & Resource economics 33: 223-250

Brander L M, van Beukering P and Cesar H (2007) The recreational value of coral reefs: a meta-analysis. Ecological Economics 63(1): 209-218

Carson R T, Flores N E, Martin K M and Wright J L (1996) Contingent valuation and revealed preference methodologies: Comparing the estimates for quasi-public goods. Land Economics 72(1): 80-99

Desvousges W H, Johnson F R and Banzhaf H S (1998) Environmental policy analysis with limited information: Principles and applications of the transfer method. Cheltenham, Edward Elgar

Engel S (2002) Benefit Function Transfer versus Meta-analysis as Policy-Making Tools: A Comparison. In: R. J. G. M. Florax, P. Nijkamp and K. G. Willis (eds) Comparative environmental economic assessment, Cheltenham, Edward Elgar

Fromm O (2000) Ecological structure and functions of biodiversity as elements of its total economic value. Environmental and Resource Economics 16(3): 303-328

Glass G V, McGaw B and Smith M L (1981) Meta-analysis in Social Research. Sage

Hoehn J P (2006) Methods to address selection effects in the meta regression and transfer of ecosystem values. Ecological Economics 60(2): 389-398

Jacobsen J B, Boiesen J H, Thorsen B J and Strange N (2008) What's in a name? The use of quantitative measures versus 'Iconised' species when valuing biodiversity. Environmental & Resource economics 39(3): 247-263

Jacobsen J B and Hanley N (2009) Are there income effects on global willingness to pay for biodiversity conservation. Environmental and Resource Economics (In press)

Johnston R J, Besedin E Y, Iovanna R, Miller C J, Wardwell R F and Ranson M H (2005) Systematic variation in willingness to pay for aquatic resource improvements and implications for benefit transfer: a meta-analysis. Canadian Journal of Agricultural Economics 53(2-3): 221-248

Johnston R J, Besedin E Y and Ranson M H (2006) Characterizing the effects of valuation methodology in function-based benefits transfer. Ecological Economics 60(2): 407-419

Johnston R J, Kukielka J B and Duke J M (2008) Systematic Variation in Willingness to Pay for Agricultural Land Preservation. Paper presented at the 2008 Workshop on Meta-Analysis in Economics and Business, Nancy, France, October 17-18, 2008.

Kristofersson D and Navrud S (2007) Can Use and Non-Use Values be Transferred Across Countries? In: Navrud S and Ready R (ed) Environmental Value Transfer: Issues and Methods, Springer, Dordrecht, The Netherlands

Kriström B and Riera P (1996) Is the income elasticity of environmental improvements less than one? Environmental and Resource Economics 7: 45-55

Lindhjem H (2007) 20 Years of stated preference valuation of non-timber benefits from Fennoscandian forests: A meta-analysis. Journal of Forest Economics 12(4): 251-277

Lindhjem H and Navrud S (2008a) How reliable are meta-analyses for international benefit transfers? Ecological Economics 66(2-3): 425-435.

Lindhjem H and Navrud S (2008b) Internet CV surveys - a cheap, fast way to get large samples of biased values? Department of Economics and Resource Management, Norwegian University of Life Sciences

Lindhjem H and Navrud S (2009) Asking for Individual or Household Willingness to Pay for Environmental Goods. Implication for aggregate welfare measures, Environmental and Resource Economics (In press)

Loomis J B (1992) The Evolution Of A More Rigorous Approach To Benefit Transfer - Benefit Function Transfer. Water Resources Research 28(3): 701-705

Loomis J B and White D S (1996) Economic benefits of rare and endangered species: Summary and meta-analysis. Ecological Economics 18(3): 197-206

Millennium Ecosystem Assessment (MEC) (2005) Synthesis Report. Washington, Island Press

Moeltner K, Boyle K and Paterson R W (2007) Meta-analysis and benefit transfer for resource valuation - addressing classical challenges with Bayesian modeling. Journal of Environmental Economics and Management 53: 250-269

Moeltner K and Rosenberger R (2008) Predicting Resource Policy Outcomes via Meta-Regression: Data Space, Model Space, and the Quest for 'Optimal Scope'. The B.E. Journal of Economic Analysis & Policy 8(1), Article 31

Nelson J P and Kennedy P E (2009) The Use (and Abuse) of Meta-Analysis in Environmental and Natural Resource Economics: An Assessment. Environmental and Resource Economics (In press)

Pearce D (2005). Economists and biodiversity conservation: What can we contribute? European Association of Environmental and Resource Economists, Bremen, 2005

Ready, R and Navrud, S (2006) International benefit transfer: Methods and validity tests. Ecological Economics 60: 429-434

Richardson L and Loomis J (2009) The total economic value of threatened, endangered and rare species: An updated meta-analysis. Ecological Economics (In press)

Rosenberger R and Johnston R J (2007). Selection effects in meta-valuation function transfers. Aarhus Colloqium of Meta-Analysis in Economics. Sønderborg, Denmark

Rosenberger R S and Loomis J B (2000a) Using meta-analysis for benefit transfer: In-sample convergent validity tests of an outdoor recreation database. Water Resources Research 36(4): 1097-1107

Rosenberger R and Loomis J (2000b) Panel stratification in meta-analysis of economic studies: an investigation of its effects in the recreation valuation literature. Journal of Agricultural and Applied Economics 32(1): 131-149

Rosenberger R and Stanley T D (2007) Publication Effects in the Recreation Use Values Literature. Annual Meeting of the American Agricultural Research Association, July 29-August 1, 2007, Portland, Oregon

Santos J M L (1998) The Economic Valuation of Landscape Change. Theory and Policies for Land Use and Conservation. Cheltenham, Edward Elgar

Shrestha R K and Loomis J B (2003) Meta-Analytic Benefit Transfer of Outdoor Recreation Economic Values: Testing Out-of-Sample Convergent Validity. Environmental & Resource economics 25: 79-100

Smith V K and Pattanayak S K (2002) Is Meta-Analysis a Noah's Ark for Non-Market Valuation? Environmental and Resource Economics 22: 271-296

Stanley T D and Jarrel S D (2005) Meta-Regression Analysis: A Quantitative Method of Literature Surveys. Journal of Economic Surveys 19(3): 299-308

Stapler R W and Johnston R J (2009) Meta-Analysis, Benefit Transfer, and Methodological Covariates: Implications for Transfer Errors. Environmental and Resource Economics (In press)

U.S. Environmental Protection Agency (2006) Report of the EPA work group on VSL meta-analysis, report NCEE-0494. National Center for Environmental Economics

Van Houtven G, Powers J and Pattanayak S K (2007) Valuing water quality improvements using meta-analysis: Is the glass half-full or half-empty for national policy analysis? Resource and Energy Economics 29(3): 206-228

Veisten K, Hoen H F, Navrud S and Strand J (2004) Scope insensitivity in contingent valuation of complex environmental amenities. Journal of Environmental Management 73(4): 317-331

Zandersen M and Tol R S J (2009) A meta-analysis of forest recreation values in Europe. Journal of Forest Economics 15(1-2): 109-130.

Appendix A: Descriptions of and full references for meta-analysis datasets:

Table A1 Studies of endangered species used in meta-analysis (MA) – Level 1

Name of references	Country	Year[a]	Species	Method	Mode	Vehicle[b]	Payment format	# of values[c]	WTP (USD)[d]
Jakobsson and Dragun (2001)	Australia	2000	Possum	CV: DC, PC	Mail	M	Year	12	14-72
Tisdell et al (2005)	Australia	2002	Glider	CV: OE	Mail	V	One-off	3	13-19
Kontoleon and Swanson (2003)	China	1998	Panda	CV: DC	Interview	M&V	Visit	3	5-17
Jianjun et al (2006)[e]	China	2005	Turtles	CV: PL	Drop-off	M	Month	5	5-6
Jianjun (2006)	China	2005	Spoonbill	CV: DC	Drop-off	M&V	Month	17	4-19
Jianjun et al (2006)[e]	Philippines	2005	Turtles	CV: DC	Drop-off	M&V	Month	5	1-3
Harder et al (2006)	Philippines	2005	Eagles	CV: DC	Drop-off	M&V	Month	30	1-3
Indab (2006)	Philippines	2005	Shark	CV: DC	Drop-off	M	Month	5	2-4
Bandara and Tisdell (2004)	Sri Lanka	2002	Elephants	CV: DC	Drop-off	V	Month	4	20-40
Bandara and Tisdell (2005)	Sri Lanka	2001	Elephants[f]	CV: DC	Interview	V	Month	6	34-41
Jianjun et al (2006)[e]	Thailand	2005	Turtles	CV: DC	Drop-off	M&V	Month	5	3-8
Nabangchang (2006)	Thailand	2005	Multiple[g]	CV: DC	Interview	M&V	One-off	7	43-64
Jianjun et al (2006)[e]	Vietnam	2005	Turtles	CV: DC	Drop-off	M	Month	4	0.2-4
Thuy (2006)	Vietnam	2005	Rhino	CV: DC	Drop-off	M&V	Month	2	13-14
Tuan et al (2008)	Cn,Pp,Th,Vn[h]	2005	Turtles	CV: DC	Drop-off	M&V	Month	16	1-5
Ninan and Sathyapalan (2005)*	India	2000	Elephants	CV:DC	Interview	M	Year	5	341-1830
Total number of studies=16								**129**	

Notes: *4 of 5 observations from this study were excluded by the screening criterion (2x STD of mean) for the Level 1 data, but included in the Level 2 dataset (see Table B)

[a] Year of data.

[b] Payment vehicle: mandatory (M) or voluntary (V).

[c] Number of WTP values used in MA.

[d] WTP values in US$. The WTP formats are given as reported (i.e. lump sum, per month, per year, per visit, per individual or household). WTP values in local currencies are converted to US$ using PPP adjustments; and values from different years are converted to 2006 prices using CPI.

[e] Jianjun et al (2006) has four separate country case study components.

[f] Abundance of elephants.

[g] Multiple species: Dugong dugong, elephants, rhinos, Irawaddy dolphin, tigers.

Table A2 Studies of nature and biodiversity conservation used in meta-analysis (MA) – Level 2

Name of references	Country	Year[a]	Habitat/service type	Method	Mode	Vehicle	Payment format	# of values	WTP (USD)
Jakobsson and Dragun (2001)	Australia	2000	Flora & fauna	CV: DC, PC	Mail	M	year	7	24-175
Bennett et al (1998)	Australia	1996	Wetlands	CV: DC	Mail	M	One-time	2	122-187
Bennett (1984)	Australia	1979	Nature reserve	CV: OE	Interview	M&V	One-time	1	33
Blamey et al (1999)	Australia	1999	Water	CA	Interview	M	Year	4	29-116
Cameron and Quiggin (1994)	Australia	1991	Parks	CV: IB	Interview	M		4	228-664
Carr and Mendelsohn (2003)	Australia	2000	Reefs	TCM	Interview		Year	1	391
Carson et al (1994)	Australia	1990	parks	CV: DBDC	Interview	M	Year	4	30-129
Hundloe (1990)	Australia	1986	Reefs	TCM	Interview		Year	1	8
Kuosmanen et al (2003)	Australia		Parks	TCM	Mail		Year	6	54-418
Lockwood and Carberry (1998)	Australia	1997	Vegetation	CM, CV	Mail	M	One-time	8	35-90
Lockwood and Tracy (1995)	Australia	1993	Parks	CV: OE	Mail	V	One-time	1	21
Lockwood (1999)	Australia	1995	Parks	CV: OE	Computer	V		4	14-450
Lockwood (1996)	Australia		Natural environm.	CV: DC	Mail	V		9	5-123
Loomis et al (1993)	Australia		Forests	CV: OE, DC	Mail		Year	6	34-89
Morrison et al (2002)	Australia	1997	Wetlands	CM	Mail	M	One-time	18	25-117
Nillesen et al (2005)	Australia		Parks	TCM	Mail		Year	1	86
Streever et al (1998)	Australia	1996	Wetlands	CV: OE	Mail	M		1	151
Greiner and Rolfe (2004)	Australia	2000	Parks	CV: OE	Interview	M	Visit	3	23-39
Campbell and Reid (2000)	Australia	1996	Fisheries	CV: DC	Interview	M	Year	3	212-517
Flatley and Bennett (1995)	Vanuatu	1994	Forest	CV	Interview	V	One-time	2	33-36
Flatley and Bennett (1996)	Vanuatu	1994	Forest	CV	Interview	V	One-time	1	18
Chen et al (2004)	China	1999	Beaches	TCM	Interview	M	Visit	1	64
Day and Mourato (2002)	China	1997	Rivers	CV: DBDC	Interview	M		4	51-94
Gong (2004)	China	2002	National reserve	CV: BG	Interview	M		2	5-16
Guo et al. (2001)	China	1997	Ecosystem services	TCM	Interview		Visit	3	20-40
Jim and Chen (2006)	China	2003	Urban green spaces	CV: PC	Interview			1	15
Yaping (1998)	China	1996	Lakes	CV:OE& TCM	Interview	M	Visit Year	7	77-114 6-15
Zhongmin et al (2003)	China	2001	Ecosystem services	CV: PC	Interview	M&V	Year	3	8-87
Zhongmin et al (2006)	China	2003	Watershed	CV:DC, DBDC	Interview	M&V	Year	2	71
Wang et al (2007)	China	2006	Water	CV:DC	Interview	M	Month	2	
Xu et al (2007)	China	2002	Eco-services	CM	Interview	M	Year	7	51-134
Gundimeda and Kathuria (2003)	India	2003	Water	HPM	Interview			2	149-377
Hadker et al (1997)	India	1995	Parks	CV: DC	Interview	V	Month	2	6-8
Kohlin (2001)	India	1995	Woodlots	CV: DC, OE	Interview		Month	11	4-6
Maharana et al (2000)	India	1998	Lakes	CV: IB	Interview		Year	4	5-43

				& TCM					
Nallathiga (2004)	India	1995	Rivers	CV: PC	Interview	M	Year	2	22-25
Butry and Pattanayak (2001)	Indonesia	1996	Forests	CV:OE, PC & MP	Interview	M	Year	3	23-2006
Pattanayak (2001)	Indonesia	1996	Ecological services	CV: DC	Interview	M	Year	1	20
Pattanayak and Kramer (2001)	Indonesia	1996	Watershed	CV: DC	Interview	M	Year	10	7-21
Walpole et al (2001)	Indonesia	1995	Parks	CV: DBDC	Interview	M	Year	1	78
Amirnejad et al (2006)	Iran	2004	Forests	CV: DC	Interview	M	Month	1	9
Fleischer and Tsur (2000)	Israel	1997	Landscapes	TCM	Interview		Year	2	179-367
Shechter et al (1998)	Israel	1993	Parks	CV: OE, DC	Telephone	V	One-time	12	28-57
Tsgue and Washida (2003)	Japan	1998	Natural areas of the Sea.	CV: DC	Internet		One-time	6	132-159
Nishizawa et al (2007)	Japan	2003	Eco-services	CVM:DC	Mail	M		2	13-14
Kwak et al (2003)	Korea	2001	Forests	CV: DC	Interview	M	Year	4	3-6
Lee (1997)	Korea	1996	Nature-based tourism resources	CV: DC	Interview	M	Year	2	12-13
Lee and Han (2002)	Korea	1999	Parks	CV: DC	Interview	M	Year	10	8-23
Lee and Mjelde (2007)	Korea	2005	Eco-services	CV:DC	Interview	M	Year	2	22-26
Eom and Larson (2006)	Korea	2000	Water	CV	Interview	M	Year	2	35-62
Lee and Chun (1999)	Korea	1994	Forest recreation	CV:DC	Mail	V	Year	3	445-787
Othman et al (2004)	Malaysia	1999	Forests	CM	Interview		Year	5	0.5-8
Yeo (2002)	Malaysia	1998	Parks	CV: OE	Interview		Year	6	6-12
Mourato (2002)	Malaysia	1997	Water	CV: PL	Interview	M	Month	2	3
Naylor and Drew (1998)	Micronesia	1996	Mangroves	CV	Interview	M	Month	4	174-556
Khan (2004)	Pakistan	2003	Parks	TCM	Interview		Visit	2	13-18
Manoka (2001)	P.N.Guinea	1999	Forests	CV: OE, DC	Mail	V	Year	10	11-101
Arin and Kramer (2002)	Philippines	1997	Marine sanctuary	CV: PC	Interview	M	Year	3	21-34
Calderon et al (2005)	Philippines	2003	Watersheds	CV: DC	Interview	M	Month	36	2-6
Choe et al (1996)	Philippines	1992	Water	CV: DC, BG, OE	Interview		Month	18	0-13
Pattanayak and Mercer (1998)	Phillippines	1994	Soil	MP	Interview		Year	2	195-306
Subade (2005)	Philippines	2002	Reefs	CV: DC	Interview & drop-off	V	Year	12	15-83
Amponin et al (2007)	Philippines	2006	Watershed	CV:DC	Interview	M	month	9	3-6
Wei-Shiuen and Robert (2005)	Singapore	2002	Beaches	TCM & CV	Interview		Year	14	0.1-485
Bogahawatte (1999)	Sri Lanka	1997	Forests	MP	Interview	Year	Year	30	1-437
Ekanayake and Abeygunawardena (1994)	Sri Lanka	1992	Forests	CV: OE	Interview		Year	2	41-131
Chang and Ying (2005)	Taiwan	2001	Agri. lands	CV: DBDC	Telephone	M	Year	1	103
Chen (1998)	Taiwan		Agri. lands	CV:OE, DC	Mail		Month	6	0.3-7
Hammit et al (2001)	Taiwan	1993	Wetlands	CV: DC, OE	Interview	V	Year	4	46-173

Cushman (2004)	Thailand	2001	Beaches	CM	Interview	M	Year	5	17-526
Isangkura (1998)	Thailand	1996	Parks	CR & CV: OE	Interview	M	Year	9	2-28
Seenprachawong (2002)	Thailand	2002	Coastal ecosystem	CM	Interview	V	Year	5	9-188
Seenprachawong (2001)	Thailand	2000	Reefs	CV:DC &TCM	Interview	V	Year	3	31-555
Tapvong and Kruavan (1999)	Thailand	1998	Rivers	CV: DC	Interview	M	Month	2	9-10
Pham and Tran (2000)	Vietnam	2000	Reefs	CV:PC & TCM	Interview	M	Year	5	7-170
Pham et al (2000)	Vietnam	2003	Reefs	TCM CV	Interview	M	Year	4	17-390
Phuong and Gopalakrishnan (2004)	Vietnam	2001	Water	CV: PC	Interview	M	Year	7	4-40
Do (2007)	Vietnam	2006	Wetlands	CM	Interview		Year	2	4-12
Nam et al (2001)	Vietnam	1999	Forests	CV & MP	Interview	M	Year	9	24-1807
Total number of observations*								**421**	

Notes: * The total number of observations using the least strict screening criterion (WTP>0 and within 2x STD of the mean), i.e. 129 (Level 1) + 421 (Level 2) = 550 observations. Blank space means that information was not reported in the study. See also notes to Table A above.

References of studies included in the meta-analysis

Amirnejad H, Khalilian S, Assareh M H and Ahmadian M (2006) Estimating the existence value of north forests of Iran by using a contingent valuation method. Ecological Economics 58(4): 665-675

Amponin J A R, Bennagen M E C, Hess S and Cruz J D S d (2007) Willingness to pay for watershed protection by domestic water users in Tuguegarao City, Philippines. PREM 07/06 Working paper

Arin T and Kramer R A (2002) Divers' willingness to pay to visit marine sanctuaries: an exploratory study. Ocean Coastal Manage. 45(2-3): 171-183

Bandara D and Tisdell C (2004) The net benefit of saving the Asian elephant: a policy and contingent valuation study. Ecological Economics 48: 93- 107

Bennett J, Morrison M and Blamey R (1998) Testing the validity of responses to contingent valuation questioning. Australian Journal of Agricultural and Resource Economics 42(2): 131-148

Bennett J W (1984) Using Direct Questioning to Value the Existence Benefits of Preserved Natural Areas Australian Journal of Agricultural Economics 28(2-3): 136-152

Blamey R, Gordon J and Chapman R (1999) Choice Modelling: Assessing the Environmental Values of Water Supply Options. Australian Journal of Agricultural and Resource Economics 43(3): 337 - 357

Bogahawatte C (1999). Forestry Policy, Non-timber Forest Products and the Rural Economy in the Wet Zone Forests in Sri Lanka. Research report, EPPSEA.

Butry D T and Pattanayak S K (2001). Economic Impacts of Tropical Forest Conservation: The Case of Logger Households around Ruteng Park. Working Paper 01_01, Environment and Natural Resource Economics Program, Research Triangle Institute, North Carolina.

Calderon M M, Camacho L D, Carandang M G, Dizon J T, Rebugio L L and Tolentino N L (2005). A Water User Fee For Households in Metro Manila, Philippines. EEPSEA research report, No. 2005-RR2.

Cameron T A and Quiggin J (1994) Estimation Using Contingent Valuation Data From A Dichotomous Choice With Follow-Up Questionnaire. J.Environ.Econ.Manage. 27(3): 218-234

Campbell H P and Reid C R M (2000) Consumption Externalities in a Commercial Fishery: The Queensland Beam Trawl Fishery. Economic Record 76(232): 1-14

Carr L and Mendelsohn R (2003) Valuing Coral Reefs: A Travel Cost Analysis of the Great Barrier Reef. Ambio 32(5): 353-357

Carson R T, Wilks L and Imber D (1994) Valuing the Preservation Of Australia Kakadu Conservation Zone. Oxf. Econ. Pap.-New Ser. 46: 727-749

Chang K and Ying Y (2005) External Benefits of Preserving Agricultural Land: Taiwan's Rice Fields. The Social Science Journal 42: 285-293

Chen W, Hong H, Liu Y, Zhang L, Hou X and Raymond M (2004) Recreation Demand and Economic Value: An Application of Travel Cost Method for Xiamen Island. China Economic Review

Chen Y (1998). Biodiversity in the Framework of Ecological Economics, University of Wisconsin Milwaukee.

Choe K A, Whittington D and Lauria D T (1996) The economic benefits of surface water quality improvements in developing countries: A case study of Davao, Philippines. Land Econ. 72(4): 519-537

Cushman C (2004). External costs from increased island visitation: case study from Southern Thailand. Working paper, Department of Resource Economics, University of Massachusetts.

Day B and Mourato S (2002) Valuing River Water Quality in China In: Pearce D, Pearce C and Palmer C (eds) Valuing the Environment in Developing Countries: Case Studies, Edward Elgar

Do N T (2007). Impacts of dykes on wetland values in Vietnam's mekong river delta: a case study in the plain of reeds. Research report submitted to EEPSEA, IDRC.

Ekanayake E R M and Abeygunawardena P (1994) Valuation of Conservation Commodity of the Sinharaja Forest: Towards Total Economic Value. Sri Lanka Journal of Agricultural Economics 2(1): 115-129

Eom Y-S and Larson D M (2006) Valuing housework time from willingness to spend time and money for environmental quality improvements. Review of the Economics of the Household 4: 205-227

Flatley G and Bennet J (1995). International Values of Tropical Forest Conservation: A Cross-cultural Contingent Valuation Experiment. Paper presented to the 39th Australian Agricultural Economic Society Annual Conference. Perth, Western Australia

Flatley G W and Bennet J W (1996) Using Contingent Valuation to Determine Australian Tourists' Values for Forest Conservation in Vanuatu. Economic Analysis and Policy 26(2): 111-127

Fleischer A and Tsur Y (2000) Measuring the recreational value of agricultural landscape. Eur Rev Agric Econ 27(3): 385-398

Gong Y (2004). Distribution of Benefits and Costs among Stakeholders of a Protected Area: An Empirical Study from China. EEPSEA Research Report 3, IDRC

Greiner R and Rolfe J (2004) Estimating consumer surplus and elasticity of demand of tourist visitation to a region in North Queensland using contingent valuation. Tourism Economics 10(3): 317-328

Gundimeda H and Kathuria V (2003) Estimation of Economic Value of Water scarcity and quality in Chennai, India: The Hedonic Approach. South Asian Network of Economic Initiatives Project

Guo Z, Xiao X, Gan Y and Zheng Y (2001) Ecosystem functions, services and their values - a case study in Xingshan County of China. Ecological Economics 38(1): 141-154

Hadker N, Sharma S, David A and Muraleedharan T R (1997) Willingness-to-pay for Borivli National Park: Evidence from a contingent valuation. Ecol. Econ. 21(2): 105-122

Hammit J K, Liu J-T and Liu J-L (2001) Contingent Valuation of a Taiwanese Wetland. Environment and Development Economics 6: 259-268

Harder D S, Labao R and Santos F I (2006). Saving the Philippine eagles: How much will this cost and are Filipinos willing to pay for it? Paper presented at the 25th EEPSEA biannual workshop. Singapore.

Hundloe T (1990) Measuring the Value of the Great Barrier Reef. Australian Parks and Recreation 26(3): 11-15

Indab A (2006). Rationalizing WTP for whale shark conservation in Sorsogon province. Paper presented at the 26th EEPSEA biannual workshop. Cebu, the Philippines.

Isangkura A (1998). Environmental Valuation: An Entrance Fee System for National Parks in Thailand. Economy and Environment Program for Southeast Asia, IDRC

Jakobsson K M and Dragun A K (2001) The Worth of a Possum: Valuing Species with the Contingent Valuation Method. Environmental and Resource Economics 21: 287-302

Jianjun J (2006). Economic valuation of the Black-faced Spoonbill conservation in Macao. Paper presented at the 25th EEPSEA biannual workshop. Singapore.

Jianjun J, Indab A, Nabangchang O, Thuy T D, Harder D and Subade R (2006). WTP for Marine Turtle Conservation: A Cross-country Comparison in Asia.. Paper presented at the Third World Congress of Environmental and Natural recourses economics, Kyoto, Japan

Jim C Y and Chen W Y (2006) Recreation - Amenity Use and Contingent Valuation of Urban Greenspaces in Guangzhou, China. Landscape and Urban Planning 75(1-2): 81-96

Khan H (2004). Demand of eco-tourism: estimating recreational benefits from the Margalla Hills national park in the northern Pakistan. SANDEE working paper, No. 5-04.

Kohlin G (2001) Contingent valuation in project planning and evaluation: the case of social forestry in Orissa, India. Environment and Development Economics 6(02): 237-258

Kontoleon A and Swanson T (2003) The Willingness to Pay for Property Rights for the Giant Panda: Can a Charismatic Species Be an Instrument for Nature Conservation? Land Economics 79(4): 483-499

Kuosmanen T, Nillesen E and Wesseler J (2003). Does Ignoring Multi-Destination Trips in the Travel Cost Method Cause a Systematic Downward Bias? Discussion Paper No. 9 2003, Mansholt Graduate School.

Kwak S J, Yoo S H and Han S Y (2003) Estimating the public's value for urban forest in the Seoul metropolitan area of Korea: A contingent valuation study. Urban Studies 40(11): 2207-2221

Lee C-K and Han S-Y (2002) Estimating the Use and Preservation Values of National Parks' Tourism Resources Using a Contingent Valuation Method. Tourism Management 23: 531-540

Lee C-K and Mjelde J W (2007) Valuation of ecotourism resources using a contingent valuation method: The case of the Korean DMZ. Ecological Economics 63: 511-520

Lee C K (1997) Valuation of nature-based tourism resources using dichotomous choice contingent valuation method. Tourism Management 18(8): 587-591

Lee H-C and Chun H-S (1999) Valuing environmental quality change on recreational hunting in Korea: A contingent valuation analysis. Journal of Environmental Management 57(1): 11-20

Lockwood M (1996) Non-compensatory preference structures in non-market valuation of natural area policy. Aust. J. Agric. Econ. 40(2): 85-101

Lockwood M (1999) Preference structures, property rights, and paired comparisons. Environmental and Resource Economics 13(1): 107-122

Lockwood M and Carberry D (1998). Stated Preference Surveys of Remnant Native Vegetation Conservation, Johnstone Centre, Albury, NSW.

Lockwood M and Tracy K (1995) Nonmarket Economic Valuation Of An Urban Recreation Park. Journal of Leisure Research 27(2): 155-167

Loomis J, Lockwood M and Delacy T (1993) Some Empirical-Evidence On Embedding Effects In Contingent Valuation Of Forest Protection. Journal of Environmental Economics and Management 25(1): 45-55

Maharana I, Rai S C and Sharma E (2000) Valuing ecotourism in a sacred lake of the Sikkim Himalaya, India. Environmental Conservation 27(3): 269-277

Manoka B (2001). Existence Value: A Re-appraisal and Cross-Cultural Comparison. Economy and Environment Program for Southeast Asia (EEPSEA) Research Report No. 2001-RR1, IDRC.

Morrison M, Bennett J, Blamey R and Louviere J (2002) Choice modeling and tests of benefit transfer. Am. J. Agr. Econ. 84(1): 161-170

Mourato S (2002) Valuing Improvements to Sanitation in Malaysia. In: D. Pearce, C. Pearce and C. Palmer (ed) Valuing the Environment in Developing Countries: Case Studies, Cheltenham, Edward Elgar

Nabangchang O (2006). Motivations for Charitable Behavior for Protection of Wildlife and Endangered Species of Thailand.

Nallathiga R and Paravastu R B (2004). Benefit Estimation of River Water Quality Conservation Using Contingent Valuation Survey: a Case-Study in the Yamuna River Sub-basin. European Association of Environmental and Resource Economics Thirteenth Annual Conference. Budapest, Hungary.

Nam M V, Nhan N T, Trihn B V and Thong P L (2001). Forest Management Systems in the Mekong River Delta, Vietnam. EEPSEA research report.

Naylor R and Drew M (1998) Policy Options: Valuing mangrove resources in Kosrae, Micronesia. Environment and Development Economics 3: 471-490

Nillesen E, Wesseler J and Cook A (2005) Estimating the Recreation-Use Value for Hiking in Bellenden Ker National Park, Australia. Environmental Management 36(2): 311-316

Ninan K N and Sathyapalan J (2005) The economics of biodiversity conservation: a study of a coffee growing region in the Western Ghats of India. Ecological Economics 55(1): 61-72

Nishizawa E, Kurokawa T and Yabe M (2007) Policies and resident's willingness to pay for restoring the ecosystem damaged by alien fish in Lake Biwa, Japan. Environmental Science & Policy 9: 448-456

Othman J, Bennet J and Blamey R (2004) Environmental values and resource management options: a choice modelling experience in Malaysia. Environment and Development Economics 9: 803–824

Pattanayak S K (2001). How Green Are These Valleys? Combining Revealed and Stated Preference Methods to Account for Ecosystem Costs of Deforestation, Environment and Natural Resource Economics Program, Research Triangle Institute

Pattanayak S K and Kramer R A (2001) Pricing ecological services: Willingness to pay for draught mitigation from watershed protection in eastern Indonesia. Water Resour. Res. 37(3): 771-778

Pattanayak S K and Mercer D E (1998) Valuing soil conservation benefits of agroforestry: contour hedgerows in the Eastern Visayas, Philippines. Agricultural Economics 18(1): 31-46

Pham K N and Tran V H S (2000). Recreation Value of the Coral-surrounded Hon Mun Islands in Vietnam. Economy and Environment Program for Southeast Asia, IDRC.

Phuong D M and Gopalakrishnan C (2004) Optimal Management of Water for Sustainable Fisheries and Aquaculture International Journal of Water Resources Development 20(4): 493-506

Seenprachawong U (2001). An Economic Analysis of Coral Reefs in the Andaman Sea of Thailand. EEPSEA research report.

Seenprachawong U (2002). An Economic Valuation of Coastal Ecosystems in Phang Nga Bay, Thailand. EEPSEA Research Report No. 2002-RR5, IDRC.

Shechter M, Reiser B and Zaitsev N (1998) Measuring Passive Use Value: Pledges, Donations and CV Responses in Connection with an Important Natural Resource. Environmental and Resource Economics 12: 457-478

Streever W, Callaghan-Perry J M, Searles A, Stevens T and Svoboda P (1998) Public attitudes and values for wetland conservation in New South Wales, Australia. Journal of Environmental Management 54: 1-14

Subade R F (2005). Valuing Biodiversity Conservation in a World Heritage Site: Citizens' Non-Use Values for Tubbataha Reefs National Marine Park, Philippines. Economy an Environment Program for Southeast Asia (EEPSEA), IDRC.

Tapvong C and Kruavan J (1999). Water Quality Improvements: A Contingent Valuation Study of the Chao Phraya River. EEPSEA, IDRC.

Thuy T D (2006). Willingness to Pay for Conservation of Vietnamese Rhinoceros. Paper presented at the 26th EEPSEA biannual workshop. Cebu, the Philippines.

Tisdell C, Wilson C and Nantha H S (2005) Policies for Saving a Rare Australian Glider: Economics and Ecology. Biological Conservation 123: 237-248

Tsuge T and Washida T (2003) Economic valuation of the Seto Inland Sea by using an Internet CV survey. Marine Pollution Bulletin 47: 230-236

Tuan T H, Thuy T D, Jianjun J, Indab A and Nabangchang O (2008). WTP for Marine Turtle Conservation in Four Countries: Benefit Transfer Testing. Final report submitted to EEPSEA, July, 2008.

Walpole M J, Goodwin H J and Ward K G R (2001) Pricing policy for tourism in protected areas: Lessons from Komodo National Park, Indonesia. Conserv. Biol. 15(1): 218-227

Wang X, Bennett J, Xie C, Zhang Z and Liang D (2007) Estimating non-market environmental benefits of the Conversion of Cropland to Forest and Grassland Program: A choice modeling approach. Ecological Economics 63(1): 114-125

Wei-Shiuen N and Robert M (2005) The impact of sea level rise on Singapore. Environment and Development Economics 1002: 201-215

Xu Z M, Cheng G D, Bennet J, Zhang Z Q, Long A H and Kunio H (2007) Choice modeling and its application to managing the Ejina Region, China. Journal of Arid Environments 69: 331-343

Yaping D (1998). The Value of Improved Water Quality for Recreation in East Lake, Wuhan, China: Application of Contingent Valuation and Travel Cost Methods. Economy and Environment Program for Southeast Asia, IDRC.

Yeo B H (2002) Valuing a Marine Park in Malaysia. In: Pearce D, Pearce C and Palmer C (eds) Valuing the Environment in Developing Countries: Case Studies, Edward Elgar,

Zhongmin X, Guodong C, Zhiqiang Z, Zhiyong S and Loomis J (2003) Applying Contingent Valuation in China to Measure the Total Economic Value of Restoring Ecosystem Services in Ejina Region. Ecological Economics 44: 345-358

Zhongmin X, Loomis J, Zhiqiang Z and Hamamura K (2006) Evaluating the performance of different willingness to pay question formats for valuing environmental restoration in rural China. Environment and Development Economics 11: 585-601

Appendix B: Reduced Models 1-4 and benefit transfer results

This appendix displays the results of reduced Models 1-4 for the Level 1 data (Table B1) and Level 2 data (Table B3), where variables not significant at the 20 percent level have been excluded. These models are then used to calculate transfer errors (Tables B2 & B4), which can be compared to results in Tables 8 and 9 in the main text.

Table B1 Meta-regression models for Level I: Endangered species studies

Variable#	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Constant	1.779^{***} (.010)	3.347^{***} (.000)	2.488^{***} (.000)	-9.867^{***} (.000)
DC	1.992^{***} (.004)		1.797^{***} (.004)	1.567^{***} (.000)
Hholdpay				1.775^{***} (.000)
Month			.785 (.101)	
Nonpara.	$-.264^{**}$ (.028)	$-.265^{**}$ (.027)	$-.300^{***}$ (.010)	$-.286^{***}$ (.009)
Interview	1.419^{***} (.004)			-1.105^{**} (.025)
Turtle			-1.347^{***} (.001)	$-.919^{***}$ (.000)
Mammal		2.038^{***} (.000)		1.678^{***} (.000)
Australia				-2.221^{***} (.000)
Philippin.			-1.143^{***} (.000)	
SE Asia				
EEPSEA				
Journal				
LnIncome				$.895^{***}$ (.000)
LnYear				2.309^{***} (.000)
Summary statistics:				
R^2: within	0.044	0.044	0.157	0.230
R^2: betw.	0.541	0.673	0.690	0.961
R^2: overa.	**0.391**	**0.438**	**0.548**	**0.810**
Sigma_u	.867	.701	.611	.248
Sigma_e	.470	.470	.444	.423
Rho	.772	.689	.654	.255
N	124	124	124	124
# studies	16	16	16	16

Note: *p < 0.10, **p < 0.05, ***p < 0.01. STATA 9.2 used. # Blank space means variable not included in regression.

Table B2 Median and mean transfer error (percent) for reduced Models 1-4, Level 1

	Model 1	**Model 2**	**Model 3**	**Model 4**
Median	69	44	68	25
Mean	108	77	103	44
N	124	124	124	123

Table B3 Meta-regression models for Level 2: Biodiversity and nature conservation

Variable#	Model 1	Model 2	Model 3	Model 4
Constant	3.672*** (.000)	4.854*** (.000)	4.798*** (.000)	(dropped)
SP	-.448* (.092)	-1.828*** (.000)	-1.779*** (.000)	-5.837*** (.000)
DC	.473** (.022)			
TCM				
Hhldpay				
Month	.549* (.066)	1.188*** (.000)	1.253*** (.000)	1.776*** (.000)
Nonpara			-.227* (.068)	-.315*** (.005)
Interview				
Turtle				-.878*** (.002)
Mammal	1.838*** (.001)	1.715*** (.000)	1.745*** (.000)	2.185*** (.000)
Marine	.963*** (.001)	.532** (.017)	.717*** (.001)	(dropped)
Wetland	-.822** (.036)	1.282*** (.001)	1.307*** (.000)	(dropped)
Species				
Terrestrial				
Nonuse				(dropped)
Oceania	.910*** (.003)	.755*** (.006)	.720*** (.006)	(dropped)
East				
Southeast	-.758*** (.004)		-.801*** (.000)	
South Asia				
EEPSEA		-.695*** (.001)		-.680*** (.010)
Journal				
LnIncome				.873*** (.000)
LnYear				
Summary statistics				
R^2 within	0.101	0.109	0.203	0.222
R^2:betwen	0.169	0.530	0.564	0.924
R^2:overall	**0.145**	**0.311**	**0.448**	**0.779**
Sigma_u	.909	.669	.710	.325
Sigma_e	1.111	.811	.583	.439
Rho	.401	.404	.596	.354
N	550	431	390	124
# studies	95	70	67	15

Note: *p < 0.10, **p < 0.05, ***p < 0.01. STATA 9.2 used. # Blank space means variable not included in regression.

Table B4 Median and mean transfer error (percent) for reduced Models 1-4, Level 2

	Model 1	Model 2	Model 3	Model 4
Median	67	58	46	26
Mean	10449	279	86	44
N	547	428	387	121

Appendix C: Model 6a, Level 2 and transfer error results

Table C1 displays a version of Model 6 for the Level 2 data in Table 7, based on Model 3 (instead of Model 5), where methodological variables have been taken out. This model is then used to calculate transfer errors (Table C2), comparable to Model 6 in Table 9 in the main text.

Table C1 Meta-regression Model 6a for Level 2 data

Variables	Model 6a
Constant	4.087*** (1.275)
SP	
DC	
TCM	
Hhldpay	
Month	
Nonpara	
Interview	
Turtle	
Mammal	
Marine	-.0203 (.429)
Wetland	.220 (.536)
Species	-.936*** (.284)
Terrestrial	-1.069** (.425)
Nonuse	-.577*** (.180)
Oceania	.639 (.476)
East	-.484 (.458)
Southeast	-.919* (.483)
South Asia	2.387*** (.724)
EEPSEA	.793* (.425)
Journal	-.006 (.383)
LnIncome	.076 (.102)
LnYear	-.036 (.229)
Summary statistics:	
R^2 within	.079
R^2:betwen	.383
R^2:overall	**.310**
Sigma_u	.897
Sigma_e	.626
Rho	.672
N	390
# studies	67

Note: *p < 0.10, **p < 0.05, ***p < 0.01. STATA 9.2 used. # Blank space means variable not included in regression.

Table C2 Median and mean transfer error for alternative version of Model 6, Table 7.

	Model 6a
Median	49
Mean	108
N	387