# The Use (and Abuse) of Meta-Analysis in Environmental and Natural Resource Economics: An Assessment

Jon P. Nelson*,[a] and Peter E. Kennedy**
*Pennsylvania State University, University Park, PA 16802
**Simon Fraser University, Burnaby, BC V5A 1S6

**Abstract**

Motivated by the 2006 report of a Work Group appointed by the Environmental Protection Agency (EPA), this paper examines the present state of meta-analysis in environmental economics and offers recommendations for its future use. To this end we summarize and assess 130 meta-analyses from 115 published and unpublished studies, covering seventeen topical categories in environmental and resource economics. First, we provide several generic meta-analysis models as reference points and discuss major estimation issues. Five econometric issues are identified as part of a complete analysis: (1) sample selection criteria; (2) basic data summary; (3) primary data heterogeneity; (4) heteroskedasticity; and (5) non-independence of multiple observations from primary studies. Second, a tabular summary is presented for the 130 meta-analyses with respect to estimation methods. Third, a narrative summary is presented for 19 meta-analyses, including the three value-of-statistical-life studies examined by the EPA Work Group and one analysis from each of sixteen other categories. Fourth, we offer a set of "best practice" guidelines for future meta-analyses in this and other areas of economics. Last, the paper comments on the use of meta-analytic methods for benefit transfers of environmental values.

**Keywords**: Meta-analysis, Environmental Valuation, Benefit Transfer, Econometric Modeling.
**JEL Classification**: C5, Q2, Q51

August 12, 2008

[a] Contact Information for Jon P. Nelson: E-mail: jpn@psu.edu; Tel.: 814-237-0157; Fax: 814-863-4775; Affiliation address: Department of Economics, 603 Kern Building, Pennsylvania State University, University Park, PA 16802; Mailing address: 642 Glenn Road, State College, PA 16803

We wish to thank H. Spencer Banzhaf, Randall S. Rosenberger, and V. Kerry Smith for helpful comments on earlier drafts. The usual caveats apply.

## 1. INTRODUCTION

Meta-analysis was first proposed by Glass (1976) as a method for the systematic quantitative summary of evidence across empirical studies. It currently enjoys widespread use in several areas, including the health sciences, psychology, education, marketing, and the social sciences. Application of meta-analysis in economics began in 1989-1990 with studies by Stanley and Jarrell (1989), Jarrell and Stanley (1990), Smith and Kaoru (1990a, 1990b), Walsh et al. (1989, 1990), and Weitzman and Kruse (1990). Several hundred analyses have been prepared in economics, with at least one-third in the area of environmental and resource economics. Increasingly, several meta-analyses on the same topic are reported. Our survey results show that one-half of 130 studies in the environmental area have been conducted since 2003, so this is an active and growing area of inquiry.

Most analyses in economics employ a technique referred to as meta-regression analysis. Briefly, the investigator collects a set of primary studies that contain a common empirical outcome, such as the long-run price elasticity of gasoline, the willingness-to-pay (WTP) for freshwater quality, and the influence of air pollution on property values. In contrast to controlled experiments in the natural sciences, the primary studies in economics employ different study designs, model specifications, and econometric techniques. In a meta-regression, the dependent variable is a common summary statistic or "effect-size," such as a regression coefficient for the price elasticity, a predicted value for the WTP, or significance level of the air pollution coefficient. One or more values of this statistic are drawn from each primary study. It is of crucial importance that this dependent variable is measuring the same economic concept across primary studies.[1] The explanatory variables in the regression include characteristics of the primary data, study design, valuation method, sample size, model specification, econometric methods, and other "quality" variables such as place and date of publication. Most regressors are specified as binary dummies. Most of these variables also are drawn from the primary studies, including unpublished studies in the "grey literature" (working papers, government reports, dissertations). In some analyses, the identity or characteristics of the primary investigators are used as regressors.

There are several possible objectives in a meta-analysis. First, the traditional objective is to provide a "combined" estimate of the effect-size. In the simplest model–the "fixed effect-size" model–a weighted-mean is calculated from the effect-size estimates in the primary studies, where the ideal weights are inverse variances of the estimates. Second, the objective may be to explain what determines the (typically) wide study-to-study variation or heterogeneity in effect-sizes. Rather than simply finding a

[1] Uniformity of effect sizes is achieved if the measure is defined in the same way across primary studies or because the meta-analyst has adjusted the effect-sizes to ensure that they measure the same concept. Uniformity is difficult to ensure, which might explain why meta-analysts are frequently delinquent in this respect (e.g., primary studies use different metrics for the same pollutant). Kerry Smith (private correspondence) believes that many practitioners do not pay sufficient to this issue and often pool inconsistent measures; see also Smith and Pattanayak (2002).

weighted mean, a meta-regression is employed to uncover variables that "explain" this heterogeneity. Based on the regression analysis, the investigator may test hypotheses or offer suggestions for improvements in the primary data, study design, and model specification or technique. Third, the objective may be to provide within-sample predicted values of the dependent variable under a particular set of conditions. For example, using a 10-point water quality index, Van Houtven et al. (2007) provide predicted values for WTP for three levels of freshwater quality. Fourth, an out-of-sample prediction can be employed as part of a benefit-transfer application of the empirical results (Bergstrom and Taylor 2006; Lindhjem and Navrud 2008; Smith and Pattanayak 2002). Indeed, U.S. Environmental Protection Agency (EPA 2000) guidelines for benefit-cost analysis characterize meta-analysis as "the most rigorous benefit transfer exercise" (p. 87). Fifth, meta-analysis can be used to summarize results of a single empirical study that has produced multiple estimates: examples are investigations of Superfund sites and micro-data sets for air pollution and housing values (Banzhaf and Smith 2007; Kiel and Williams 2007; Messer et al. 2006). Finally, other applications may be presented such as a study of "publication bias."[2]

Our objective in the present paper is to assess the use of meta-analysis in environmental and resource economics. There are several earlier surveys that broadly share this objective, but given the rapid growth of meta-analysis these surveys are either dated, less technical or comprehensive, or not focused sufficiently on environmental issues (Desvousges et al. 1998; Florax et al. 2002; Smith and Pattanayak 2002; Stanley 2001; Stanley and Jarrell 1989; van den Bergh et al. 1997). In particular, none of the earlier surveys examines more than a handful of prior environmental meta-analyses. Further, meta-analysts have employed a wide variety of econometric techniques and methods, and recent studies tend to use more advanced methods. Nevertheless, new analyses appear that ignore several basic methodological and econometric issues. We focus on five problems or procedures that help define a complete meta-analysis: (1) sample selection criteria; (2) basic data summary; (3) primary data heterogeneity; (4) treatment of heteroskedasticity; and (5) non-independence of multiple observations from the same primary study. These developments and problems are documented using a tabular summary of 130 meta-analyses drawn from 115 studies. In order to further illustrate major issues, a narrative summary of 19 meta-analyses is presented that covers seventeen topical categories in environmental and resource economics. As part of our assessment, a set of "best-practice guidelines" is offered for future meta-analyses in this and other areas of economics.

[2] Publication bias (aka "file-drawer problem") is a form of sample selection bias that arises if primary studies with statistically weak, insignificant, or unusual results tend not to be submitted for publication or are less likely to be published. This area has developed its own specialized set of analytical techniques; see Florax (2002b), Roberts and Stanley (2005), Rosenberger and Stanley (2007), Rothstein et al. (2005), Stanley (2005, 2008), and Sutton et al. (2000). In this paper, we regard publication bias as an issue for future research as most studies in our sample treat this bias in a cursory fashion. An exception is Smith and Huang (1993, 1995).

The present paper is motivated by the report of a Work Group appointed by EPA (2006) that was critical of three meta-regression analyses of the value of a statistical life (VSL). The EPA Work Group highlighted the high degree of heterogeneity in the primary VSL estimates, the dependencies stemming from the inclusion of multiple estimates derived from the same primary data, and the failure of the studies to adequately document their data collection criteria and procedures. The Work Group suggested a number of improvements in methodology that might help alleviate these and other problems, including better literature search and data coding protocols; weighting of each estimate according to its variance; presentation of weighted means and confidence intervals; graphical analysis of data and residuals; separate analyses of hedonic wage and contingent valuation studies; and using only one primary estimate per study. In this paper, we show that the research problems that underlie the Work Group's concerns are wide-spread in the meta-analysis literature in environmental economics.[3] We suggest improvements in methodology and exposit the circumstances in which recent econometric advances are appropriate, something not clear in the Work Group's report or the general literature. In summary, the present paper extends the Work Group's contribution in two dimensions, first, by extending its method, and second, by extending its range to other topics, including meta-analyses for economic studies of air pollution, endangered species, energy demand, global warming, hazardous waste, recreation activities, water demand and quality, wetlands valuation, environmental regulations, and preference valuation methods.

The remainder of the paper is organized into five sections. Section 2 exposits the circumstances in which several of the popular meta-analysis models employed in the literature and their associated estimation procedures are appropriate for primary data in economics. Several key concepts are discussed, including fixed-, random-, and mixed-effect-size models. We identify several major econometric problems likely to be encountered by applied researchers conducting a meta-analysis, including unobserved heterogeneity, heteroskedasticity, non-independence of multiple estimates from primary studies, and outlier problems. Section 3 contains a tabular summary of 130 meta-analyses in environmental and resource economics. The studies are grouped into seventeen topical categories. This section identifies how these analyses dealt with five issues and documents the suitability of the applied methods currently used by economists for meta-analysis. Section 4 presents a narrative summary of 19 meta-analyses, including the three VSL studies examined by the EPA Work Group and one analysis from each of sixteen other environmental categories. The goal in this section is to provide a more detailed assessment of the use of meta-analysis by economists. Section 5 presents our best-practice guidelines and Section 6 contains the conclusions, including the implications for benefit transfers using meta-analysis.

---

[3] The EPA Work Group's report was augmented by a Science Advisory Board report (EPA 2007) for the National Center for Environmental Economics, which also was critical of the current use of VSL meta-analysis.

## 2. META-ANALYSIS DATA, MODELING, AND ESTIMATION

Consider a common econometric problem such as estimation of the price elasticity of demand for gasoline. A meta-analysis seeks to combine estimates from different primary studies of this outcome and to explain the reasons behind variation in the estimates. The econometric procedures appropriate for these tasks depend on the nature of the data available for the analysis and on assumptions made by the researcher regarding the data-generating process. In general, specification of the meta-analysis model is determined by the available data, although economic theory offers some guidance or suggests useful tests and procedures (e.g., positive price elasticities can be treated as outliers). This section surveys the leading models employed in meta-analysis with the objective of clarifying the context in which statistical and econometric procedures popular in the literature are suitable.[4]

### 2.1 Data Characteristics

As a prelude to this survey, we describe three dimensions of the data that have substantive implications for the choice of a meta-analysis model.

**1. Sample Heterogeneity.** Heterogeneity refers to effect-size estimates from primary studies not all estimating the same population effect, which is surely the case in most economic studies. Christensen (2003, p. 10) identifies two basic causes of heterogeneity, factual and methodological. Factual heterogeneity is when there are real differences in effects between primary studies. For example, the gasoline price elasticity can be different in the short-run compared to the long-run or different in Europe compared to the U.S. or different in California compared to New Jersey. Factual heterogeneity tends to be exacerbated by analysts who select "comprehensive" samples. Methodological heterogeneity arises from the use of different primary study designs and methods, an outcome that is exacerbated by the publication polices of academic journals (Smith and Pattanayak 2002). For example, some primary studies might include different explanatory variables or use a different functional form and estimation method for the coefficients and standard errors. Low-quality primary studies are likely to be more heterogeneous, which provides a justification for omission of studies with particularly poor data or methodologies.

Heterogeneity is handled in two main ways in the literature. The first is via a meta-regression in which the variation is explained by regressors for study descriptors, typically binary dummies, representing observed sources of heterogeneity. This is by far the most common method in economics. The regressors are divided into several categories, such as characteristics of the environmental "commodity" examined in the primary studies, characteristics of the primary study methods and model

[4] Topics in this section are covered formally in Christensen (2003), Cooper and Hedges (1994), Hedges and Olkin (1985), Hunter and Schmidt (2004), Lipsey and Wilson (2001), Schulze (2004), and Sutton et al. (2000). Some of the technical exposition in this section follows Hox and de Leeuw (2003).

specification, and context variables such as income, location, and time period. Given uncertainty about the relevant variables, many meta-analyses present regression estimates for both full and restricted sets of explanatory variables (e.g., Loomis and White 1996; Rosenberger and Loomis 2000a; van Kooten et al. 2007). Sample size permitting, we recommend that meta-regressions also are estimated on more homogeneous subsamples, especially if a policy application is involved (e.g., de Blaeij et al 2003; Smith and Pattanayak 2002; Walsh 1989). The second method is modeling of primary study estimates as random draws from a distribution, so that each primary study (or group of studies) is estimating a different population effect-size. In the meta-analysis literature, this is known as the random-effect-size (RES) model. In contrast, the fixed-effect-size (FES) model applies when the primary studies are estimating a common, or fixed, population effect. Several formal tests for sample homogeneity are available, such as Cochrane's Q-statistic (Hedges 1994).[5] Alternative procedures simply involve testing if the coefficients of explanatory variables are significantly different from zero in a meta-regression.[6] The advantage is that coefficient signs and magnitudes also are revealed.

**2. Heteroskedasticity of Effect-Size Variances.** Due to different primary sample sizes, different sample observations and different estimation procedures, effect-size estimates generally have non-homogeneous variances. Effect estimates with smaller variances are more reliable. When these variances are "known," the analysis can account for heteroskedasticity by giving the more reliable estimates greater weight in the meta-regression.[7] However, the variance estimates of the effect-sizes may be missing for two main reasons. First, the primary studies may not report variance estimates. Second, the effect-sizes may be estimated by the meta-analyst based on empirical results in the primary studies, rather than being directly available in the studies. In this case, calculating the relevant variance may be impossible because it requires knowledge of covariances not provided in the primary studies. Access to the raw data used in the primary studies is required to solve this problem, which is rarely possible. In order to deal with missing variances, researchers have employed several alternative measurement methods. For example,

[5] The Q-statistic, distributed as a chi-square, is based on the ratio of between-study to within-study variances; Q-tests for homogeneity are reported by Brons et al. (2007), Daniel et al. (2007), de Blaeij et al. (2003), Debrezion et al. (2007), Nelson (2004), Oltmer et al. (2000), and Ringquist (2005). An alternative statistic is the Birge Ratio; see Higgins and Thompson 2002.

[6] A version of this is often referred to in the meta-analysis literature as an ANOVA-type model. This is simply an F-test of the significance of slope coefficients in a regression with one or more dummy variables as explanatory variables. For examples of ANOVA models in economic studies, see Dalhuisen et al. (2003), de Blaeij et al. (2003), Mulatu et al. (2003), Oltmer et al. (2000), Patuelli et al. (2005), and Won et al. (2007).

[7] Variances are not known, but in this literature the estimated variances from the primary studies are treated as known values. Due to the non-experimental nature of economic studies, Smith and Karou (1990b) express doubts about the use of weights in a meta-regression. They argue that "weighting implicitly assumes that the estimates based on incorrect modeling assumptions remain unbiased but simply have less informational content" (p. 425). Instead of weighting observations, some analyses have used the sample size as a regressor (e.g., Smith and Huang 1993). We discuss this bias issue below in Section 4.4.

Cavlovic et al. (2000) and Daniel et al. (2007) employ the “delta method” described in Greene (2008) to approximate variances that involve different effect-size estimates, while Mrozek and Taylor (2002) construct weights using related t-statistics. Another common method is to proxy the variances using the primary study sample sizes. Studies in environmental economics that employ this approach include Brons et al. (2007), Day (1999), de Blaeij et al. (2003), Florax et al. (2005), Horowitz and McConnell (2002), Santos (1998), Van Houtven et al. (2007), and van Kooten et al. (2007). However, our survey indicates that 95 of 130 meta-analyses did not incorporate knowledge of variances or sample sizes. Given the wide range of outcomes in primary estimates, we recommend that future researchers place a high priority on collecting primary data on variances and sample sizes.

**3. Correlation Within and Between Primary Studies.** There are several possible reasons why primary estimates of effect-sizes may not be independent of one another: (1) some primary studies may utilize the same data sources such as aggregate time-series data or public surveys; (2) analysts frequently draw more than one effect-size estimate from each primary study; (3) the analyst may apply a similar adjustment to the primary data to produce common effect-sizes; (4) several primary studies may share an unobservable characteristic such as similar management of an environmental commodity at different locations or anchoring by CV survey respondents;[8] and (5) several primary studies may share an observable characteristic, such as an identical functional form, omission of a key explanatory variable, or data drawn from the same study location. The first and last reasons can be dealt with by including the observable as a regressor, such as dummy variables for data source and location. The other three give rise to correlated errors among some group or cluster of estimates. The most common problem is the use of multiple estimates from the same primary study, so there is within-study autocorrelation. In other cases, subsets of primary studies may use the same data or embody similar specifications, leading to between-study correlated observations or data clusters.[9] Correlated effect-size estimates imply biased standard error estimates. An extreme case of interdependence is found in Florax et al. (2005), which is a meta-analysis of WTP for reductions in pesticide risk exposure. Their analysis uses only 15 primary studies, which yield a total of 331 observations of the WTP: five studies yield 1-6 estimates and five studies yield 10-18 estimates. The remaining five studies yield 24, 26, 32, 48, and 115 estimates. As this study illustrates, some means of adjusting for non-independence should be undertaken in a meta-analysis, such as use of a single estimate per primary study (or study-level averages), random selection, panel-data

[8] “Anchoring” occurs if subsequent CV bids by survey respondents are “framed” by presentation of an initial bid, which is inconsistent with accurate revelation of preferences for an environmental commodity. In a meta-analysis, anchoring leads to within-study correlated effect-sizes. We are indebted to Spencer Banzhaf for this insight.

[9] It is perhaps obvious that duplicate studies by the same author or minor extensions of an existing primary study lead to non-independent effect-size estimates. This case is discussed in depth in Wood (2008).

methods, and other econometric methods for dealing with correlated data.[10] The simplest method involves using only one estimate from each primary study, which is often recommended (Lipsey and Wilson 2001, p. 101), but this might result in an unacceptably small sample for the meta-analysis.[11] More sophisticated panel-data procedures are discussed below. Our survey indicates that 94 of 130 studies use more than one estimate from each primary study and 47 of these do not contain controls for non-independence. Another of our recommendations is that future researchers should place a high priority on adjusting for correlated effect-size estimates, within each study or across groups of studies.

In sum, data heterogeneity, heteroskedasticity, and non-independence were recognized as problems by several economists who prepared early meta-analyses (e.g., Stanley and Jarrell 1989; Smith and Kaoru 1990a, 1990b). Recent studies have addressed these issues by using a variety of econometric methods, including weighted least-squares, multilevel models, and panel-data methods. We document these techniques below using a summary of 130 meta-analyses drawn from seventeen environmental categories. Nevertheless, some recent analyses fail to deal adequately with these issues and might, for example, rely on simple OLS regressions on pooled data from diverse primary studies.

### 2.2. Generic Meta-Analysis Models

The choice of a meta-analysis model and estimation procedure depend on the available data, which is partly an artifact of the primary studies and partly determined by the analyst. Consider the example of a meta-analysis for the price elasticity of demand for gasoline. Differences in elasticity estimates could arise solely due to sampling-estimation errors within each study, in which case there is no heterogeneity. Alternatively, there could be heterogeneity because price elasticities differ from region to region, because not all studies use the same specification or estimation method, or because multiple estimates are drawn by the analyst from each primary study. Alternative meta-analysis models and their associated features can be classified on the basis of the three data features discussed above: data heterogeneity, heteroskedasticity, and non-independence. We also discuss an additional feature of meta-analytical data sets, the existence of outlier observations.[12]

---

[10] Florax (2002a) proposes the use of Moran's I and Moran's scatter-plot as devices for testing and visualizing within-study and between-study dependence. The I-test is based on the ratio of the covariance of effect-size estimates in clusters of observations to the overall variance of effect-size estimates. Florax concludes that between-study dependence is accounted for by study-descriptor regressors, but within-study dependence is not.

[11] The use of a single observation per primary study has been criticized for being ad hoc, but this procedure resolves the dependence problem and avoids using results from numerous ad hoc primary model specifications by putting zero weight on these observations. The use of a single estimate is a form of "best evidence synthesis."

[12] Our discussion focuses on estimation of the effect-size, but in some applications an interval estimate or hypothesis test is involved, which requires unbiased estimates of the variance as well. Failure to correct for heteroskedasticity or for correlated observations results in biased estimates of the parameter variances.

**1. No Heterogeneity.** Suppose that every primary study produces a single unbiased estimate of the same unknown elasticity value, and each study has been conducted in a similar fashion so the design features of the study do not affect the expected value of the elasticity estimates. Further, the estimates are stochastically independent of each other. For example, regional data on gasoline consumption might be used to estimate a long-run price elasticity of demand for the U.S., which is the same across all regions. A general model for this case is

$$\tilde{\beta}_i = \beta_i + e_i \tag{1}$$

where $\tilde{\beta}_i$ is the estimate in primary study $i$ ($i = 1, \ldots, N$), $\beta_i$ is the population value of this estimate, and $e_i$ is a sampling-estimation error. It is assumed that the sampling error is distributed normally with mean zero and variance $\sigma_i^2$. In the simple FES model, the estimates share a common, or fixed, population value, such that $\beta_1 = \beta_2 = \ldots = \beta_N = \beta$. If the variances of the primary estimates are not known and cannot be approximated, a meta-analysis is forced to give equal weight to each primary observation. This requires simply taking an unweighted average of the estimates, which is equivalent to an OLS meta-regression in which the intercept is the only regressor. When the variances are "known" (because the primary study variances $s_i^2$ are available) or can be approximated, a FES weighted-mean can be calculated, with weights the inverses of the variances; see Hedges and Olkin (1985, p. 110). This is equivalent to estimating a meta-regression by weighted least-squares, with only the intercept as a regressor. Primary study sample sizes are an alternative weighting series.

The case of "no heterogeneity" is of relatively little interest in economics because heterogeneity is undoubtedly present in virtually all data sets used for meta-analysis. However, these averages can be useful as preliminary estimators to familiarize researchers with their data or as a means of detecting outliers. Because weighted-means are important as reference points, this exercise is useful prior to conducting a meta-regression analysis. Hence, we recommend their estimation wherever feasible. Previous analyses in environmental economics that calculate FES weighted-means are Daniel et al. (2007), Debrezion et al. (2007), and Nelson (2004). This preliminary step also forces the analyst to address a basic research question: Are there good reasons to assume a universal effect-size that is common to all primary studies in a given sample? If the answer is no, this should be accounted for in the estimation procedures and applications of the empirical results.

**2. Unexplainable Heterogeneity.** It may be that there is reason to believe that there is heterogeneity in the data, either *a priori* or on the basis of a Q-test, but this heterogeneity is not explainable by using regressors. This view is frequently justified due to a large number of possible artifacts in the primary studies, which are largely unobservable by the meta-analyst. For example, gasoline price elasticities might vary due to unobserved behavior as heterogeneous consumers encounter a range of fuel economy choices and a distribution of gasoline prices (Brons et al. 2007). In this case, each effect-size observation is modeled as a random draw from a distribution with an unknown mean and unknown variance, typically a normal distribution.[13] Formally, the heterogeneity is modeled as $\beta_i = \alpha_0 + u_i$, where the mean outcome across studies is given by $\alpha_0$ and $u_i$ is an error term with mean zero and variance $\tau^2$. Adding this to equation (1) yields the random-effect-size (RES) model

$$\tilde{\beta}_i = \alpha_0 + u_i + e_i\,, \tag{2}$$

where $u_i$ and $e_i$ are assumed to be independent (Hedges 1992, p. 290), and where the concept of the true effect-size as random is distinguished from an effect estimate as random (Raudenbush 1994). By comparing the dispersion of the effect-sizes to the dispersion that would be expected given knowledge of the $\sigma_i^2$, the value of the between-study variance $\tau^2$ can be estimated. The RES variance of each primary observation is given by $v_i^2 = \sigma_i^2 + \tau^2$. Using the inverses of these new variances produces the RES weighted-mean (Hedges and Olkin 1985; Sutton et al. 2000). Alternatively, this can be viewed as estimation via weighted least-squares, with weights given by the inverse of $v_i^2$.

The case of "unexplainable heterogeneity" is rarely encountered in economics because several possible sources of heterogeneity are usually apparent and testable. Nevertheless, a few economic studies have computed RES weighted-means; see Bellavance et al. (2007), Daniel et al. (2007), Kochi et al. (2006), Nelson (2004), and Ringquist (2005). In practice, FES and RES means do not always differ substantially, but the RES model produces a larger confidence interval and gives somewhat greater weight to primary estimates with larger sampling-estimation errors relative to the FES model. It also is useful for motivating the mixed-effect-size model described below.

**3. Explainable Heterogeneity.** In this case regressors are used to "explain" why the estimates vary across primary studies. In our gasoline elasticity example, we could specify that the elasticity is different in the long-run compared to the short-run or different across geographic regions or different if estimated using a linear or double-log functional form. We could also note that some primary studies use state-level data or that estimation is based on an error components model, and so forth. This is formalized by writing the expression $\beta = \alpha_0 + \alpha_1 x_{i1} + ... + \alpha_K x_{iK}$, where $(x_{i1},...,x_{iK})$ is a vector of study descriptors, $(\alpha_1,...,\alpha_K)$ are unknown parameters (K < N), and $\alpha_0$ is the intercept term. Assuming that $\beta = \beta_i$, inserting this expression in equation (1) yields

[13] Due to the presence of skewed data, Espey and Espey (2004) use a gamma distribution for price and income elasticity estimates for electricity demand. Many other studies use a logarithmic transformation of the effect-size.

$$\tilde{\beta}_i = \alpha_0 + \alpha_1 x_{i1} + ... + \alpha_K x_{iK} + e_i \text{, where } e_i = \tilde{\beta}_i - \beta \text{ for } i = 1,...,N \tag{3}$$

Equation (2) is referred to as the "fixed-effect-size" (FES) meta-regression model. To correct for heteroskedasticity, the parameters in equation (3) are estimated by weighted least-squares with inverse variances from the primary studies ($1/s_i^2$) as analytical weights.[14] Alternatively, some studies use the sample sizes as substitute weights (e.g., Day 1999) or include the sample size (N) as a regressor.

A number of meta-analyses have employed weighted least-squares, but many do not employ inverse variance weights. Some analyses employ weights based on the number of observations obtained from each primary study, so the within-study weights sum to one (e.g., Johnston et al. 2006; Koetse et al. 2007b; Mrozek and Taylor 2002). This is intended to address an issue of sampling bias (Poe et al. 2001), which we discuss below. If primary study variances are unavailable, or thought to be poorly estimated, a heteroskedastic-consistent estimator of the meta-analysis variances should be considered, such as the White or Huber-White estimators. Finally, because most study descriptors are specified as binary dummies, the intercept term in equation (3) has a convenient interpretation as the expected effect-size for the null case (e.g., Koetse et al. 2007b; Nelson 2004; Patuelli et al. 2005; Ringquist 2005; Waters 1996).[15]

**4. Partially Explainable Heterogeneity.** It is not reasonable to expect that a meta-regression can explain all of the variation present in the data, either due to unobservables or because the estimates are drawn from a distribution of population effects. Our survey of 130 meta-analyses shows that the median adjusted R-square is 0.45, so the explanatory power of most meta-regressions is far from "perfect." This expectation can be modeled by assuming that the variation in population effects follows a linear regression model given by (Hox and de Leeuw 2003, p. 93)

$$\beta_i = \alpha_0 + \alpha_1 x_{i1} + ... + \alpha_K x_{iK} + u_i \tag{4}$$

where $u_i$ is an error term assumed to be normally distributed with zero mean and variance $\tau^2$. Substitution of equation (4) in equation (1) yields the following multilevel or mixed-effect-size (MES) regression model (Sutton et al. 2000, p. 97)

[14] These weights are the source of some confusion in the literature as some investigators refer to weights based on variances and others refer to weights based on standard errors. Some software packages require the "analytical weights," which in this case are the variances, e.g., Stata option [w] in the regress command (StataCorp 2007, p. 93). However, other software packages, such as EViews, require the user to enter a "program-weight" series, whose values are proportional to the reciprocals of the error standard deviations (QMS 2007, p. 34). These weights are used to transform the data, so regressing on the transformed data converts the standard error weights to variances. Saxonhouse (1976) first proposed using inverse standard errors as weights for a within-study meta-regression on hospital cost-function parameters; see also Greene (2008, p. 231).

[15] A number of analyses use a log transformation for the dependent variable. The coefficient estimates for dummies are interpreted as proportional changes, but this requires a transformation of the estimated coefficient that is not always handled correctly in this literature; see Kennedy (1981) and Van Garderen and Shah (2002) for guidance.

$$\tilde{\beta}_i = \alpha_0 + \alpha_1 x_{i1} + ... + \alpha_K x_{iK} + u_i + e_i \tag{5}$$

where assuming $u_i$ and $e_i$ are independent, the variance of the composite error is $v_i^2 = \sigma_i^2 + \tau^2$. The MES model can be estimated by generalized least-squares using the inverse of $v_i^2$ as weights. An estimate of $\tau^2$ can be obtained by exploiting the known $\sigma_i^2$ values, using either an iterative maximum likelihood procedure (Sutton et al. 2000, p. 98) or a non-iterative moments-estimator (Raudenbush 1994, p. 310).

According to Raudenbush (1994, p. 316), the advantages of the MES model are, first, it accounts for the extra uncertainty associated with the relevant population and, second, it permits a parsimonious summary of results when the number of studies is large. In addition, estimation of the MES model allows a determination of how much of the total variation can be explained and how much is truly random; if the latter is large, it calls into question the usefulness of a meta-analysis. In the environmental economics literature, results for the MES regression model are reported in Bellavance et al. (2007), Daniel et al. (2007), Koetse et al. (2007b), and Ringquist (2005). It is also possible to estimate in this context the more general hierarchical/multilevel model, which is discussed in the next section.

### 2.3. Correlated Effect-Size Estimates

The previous models assumed that each primary study produced only one estimate of the effect-size. However, due to specification searches and sensitivity analyses, most primary studies in economics produce more than one econometric estimate of the effect-size. Hence, the analyst is faced with the problem of using some or all of the available information. Our survey of 130 meta-analyses reveals that the median study uses three observations per primary study (mean is 6.5). Estimates from the same study are likely to be correlated, something that was recognized very early in the meta-regression literature (Stanley and Jarrell 1989, p. 166). Further, as discussed above, the primary studies may not be independent of each other. Although methodological discussions of meta-analysis make frequent mention of correlated effect-sizes, the general literature is remarkably silent on solutions to this problem.

It is useful to review why this is a problem. Suppose we have two studies, Study A reporting one effect-size estimate and Study B reporting two effect-size estimates. Suppose further that all three estimates are unbiased, but that the two coming from the Study B are highly correlated because they differ only in that they come from regressions using the same basic data but including different orthogonal sets of explanatory variables. Averaging the three estimates produces an unbiased estimate of the true effect-size. But this treats each of the three estimates as making an equal contribution to the meta-analysis, which ignores the fact that the two estimates from Study B are essentially the same. Rather than weighting each of the three estimates by one-third, we should be employing an averaging procedure which recognizes that the Study B estimates are basically contributing only one estimate and adjust the weights accordingly. If the Study B estimates were perfectly correlated, for example, the weights should be one-half on the Study A estimate and one-quarter on each of the Study B estimates. This would produce a more sensible unbiased estimate. This is the essence of what using generalized least-squares (GLS) rather than OLS accomplishes.[16] We recommend that meta-analysts consider employing GLS whenever they have reason to believe that effect-size estimates are correlated. Computational difficulties associated with doing this in the meta-analysis context require that practitioners exploit available software. Fortunately, by viewing estimates from the same primary study as hierarchical groups, panels, or clusters, estimation procedures and associated software from the hierarchical, panel-data, and clustering literature can be employed.

Hierarchical, or multilevel, regression is the procedure most closely matching meta-analysis in this respect, especially the random- and mixed-effect-size models. Hierarchical regression allows the regression coefficients to vary randomly across groups, creating composite errors in the regression that can be quite complicated; however, estimation with its related software adjusts for complex error correlation created by a composite error. The potential for allowing the slopes to vary randomly across groups makes this modeling approach very flexible. Its most common application, however, is one in which the intercept in the regression is modeled as random, but the slopes are not. This creates what was described earlier as the mixed-effect-size model, but in a context in which effect-size estimates are correlated within groups. This correlation arises because within each group all effect-size estimates embody the same random intercept; this is a special type of correlation, common to all of the estimation procedures used in this context. Each group could be a primary study or it could be a set of studies sharing some important characteristic used as the grouping criterion.[17] Examples of multilevel regression models for meta-analysis are found in Bateman and Jones (2003), Brouwer et al. (1999), Ghermandi et al. (2007), Johnson et al. (1997), and Johnston et al. (2003, 2005).[18]

[16] GLS also accomplishes a second service that should not be overlooked, namely it produces appropriate standard error estimates, so that hypothesis testing is legitimate. Although OLS may be unbiased, in the presence of correlated effect-size estimates, its traditional standard error estimates are downward biased. Moulton (1990) illustrated this dramatically. We recommend that meta-analysts should never conduct hypothesis testing with OLS estimates unless robust standard error estimates have been calculated.

[17] Rosenberger and Loomis (2000b) point out that there are several possible ways to stratify the primary data, including panel groups for each study, groups based on primary researcher, or groups based on data structure. It is not necessary that the groups, panels, or clusters be identified with individual studies.

[18] For additional discussion of multilevel models, including available software, see Bickel (2007), Cameron and Trivedi (2005), Gelman and Hill (2007), Greene (2008), Hox and de Leeuw (2003), Rabe-Hesketh and Skrondal (2008), and Raudenbush and Bryk (2002). See Johnston et al. (2003) for a meta-analysis that compares OLS, GLS, multilevel, and simultaneous equation (2SLS) models.

Formally, let $\tilde{\beta}_{ij}$ represent the *j*-th estimate obtained from the *i*-th primary study, with $J = \sum_i J_i$ total observations across studies. The hierarchical/multilevel model is written as

$$\tilde{\beta}_{ij} = \alpha_0 + \alpha_1 x_{ij1} + ... + \alpha_K x_{ijK} + u_i + e_{ij} \tag{6}$$

where the $i$ subscript now denotes the *i*-th group/cluster/study and the appearance of subscript *j* is the only substantive difference between equations (5) and (6). The random intercept for the *i*-th study is given by $\alpha_0 + u_i$. The major drawback to this procedure is that it assumes that the heterogeneity, as represented by the random intercepts, is uncorrelated with the explanatory variables. However, such correlation is a common occurrence in economics.[19] Suppose there is a group of studies with a relatively high random intercept because, for example, they use an unreported or unobserved contingent valuation (CV) survey procedure, and these studies also tend to use a particular functional form. Then the meta-analysis will attribute this high random intercept to the functional form choice and so produce a biased estimate of the slope of the functional-form dummy in the meta-regression. Further, this bias can spill over into other parameter estimates if the functional-form dummy is correlated with other explanatory variables.

There are two main ways of dealing with the correlation problem. One is to ignore it on the grounds that it is not important for purposes of the meta-analysis. For example, if the meta-analysis is to be used to forecast an effect-size, the bias in the coefficient estimates can be an advantage because it corrects forecasts for the influence of unmeasured correlated explanatory variables. As a second example, we may not be interested in the coefficients that are likely to be biased. And as a third example, we may believe that any bias created is modest so that using the mean square error criterion the parameter estimates are acceptable. We recommend that whenever a meta-analyst believes that the heterogeneity is not correlated with the explanatory variables, or that such correlation does not create much damage, the hierarchical/multilevel regression software is a good option. However, if this is not the case, a second way of dealing with the correlation problem is to use panel-data estimation procedures.

A panel data approach offers an alternative estimation procedure with both advantages and disadvantages. By interpreting each study (or each "grouping" of observations) as providing a panel of observations, panel data software can be used for estimation. Because the number of observations obtained from each grouping is unlikely to be uniform, the primary data in a meta-regression form an unbalanced panel and such panels are inherently heteroskedastic (Baltagi 2005; Rosenberger and Loomis

[19] This uncorrelatedness assumption is not always made clear in meta-analysis studies that use mixed or multilevel models, e.g., Bellavance et al. (2007). The multilevel model is especially popular outside of economics, and typically uncorrelatedness is assumed without employing any testing procedures for independence. For example, Raudenbush and Bryk (2002), a comprehensive book on the subject, do not point this assumption out until page 254, while Bickel (2007), a multilevel book written explicitly for practitioners, never mentions it at all.

2000b). In the fixed-effects (FE) panel data model the intercepts for each panel are viewed as fixed parameters, whereas in the random-effects (RE) panel data model these intercepts are viewed as random draws from a distribution.[20] The RE-model for panel data matches the mixed-effect-size model above; its estimation via panel data software produces results that are virtually identical to its estimation using hierarchical/multilevel software. The panel data approach has several advantages.

(a) For econometricians, panel data modeling and software are familiar.

(b) Panel data modeling lays stress on testing for correlation between the heterogeneity and the explanatory variables, via Hausman tests, based on the difference between FE- and RE-model estimates. Because of this, researchers using this approach are less likely to overlook testing for the correlation problem, what should be a crucial step in the modeling process.

(c) Panel data software can be used to estimate a FE-model, not available in the multilevel software. By treating the intercepts as fixed parameters, the FE-model avoids the bias caused by correlation between the heterogeneity and the explanatory variables. Further, panel data estimation has inspired two other means of dealing with this bias:

i. Use an instrumental variable estimator. A well-known example is to use one or more panel-level means and deviations from these means as instruments for explanatory variables correlated with the random intercepts; see Baltagi (2005, p. 126); Cameron and Trivedi (2005, p. 719); and Rabe-Hesketh and Skrondal (2008, p. 115). This extension, due to Hausman and Taylor (1981) and Breusch et al. (1989), does not appear to have been employed in the meta-analysis literature.

ii. Use the "between-groups" estimator, i.e., OLS where each observation is the average of data in a group as illustrated in Jeppesen et al. (2002). This approach also alleviates measurement error. Its disadvantage is that it ignores information contained in the within-group variation of the explanatory variables.[21]

Unfortunately, the FE panel data model, the main reason for using panel data software rather than hierarchical/multilevel software, has several disadvantages.[22]

---

[20] Note that the fixed- and random-effects terminology used in the panel-data literature does not match its use in the meta-analysis literature or in the hierarchical/multilevel literature.

[21] However, disregarding information contained in the within-group variation could be an advantage in the presence of anchoring in CV studies (see n.8). There are two kinds of variation in the primary studies typically used in a meta-analysis, variation of effect-sizes within a study and variation across studies. In the RE panel model, the slope estimate of interest is weighted average of within-study variation and between-study variation. In the FE panel model, the between-study variation is ignored. Hence, FE estimation is problematic in the presence of anchoring.

[22] Another disadvantage of the panel data software approach is that a variant of the usual RE panel estimation procedure is required to deal with the known heteroskedasticity; see Baltagi (2005) and Greene (2008). Tests for group-wise heteroskedasticity are found in Brown and Forsythe (1992).

(a) Because each panel has an intercept parameter, many degrees of freedom are lost by this procedure, resulting in less efficient estimates.

(b) The slope of any explanatory variable that does not vary within a panel cannot be estimated. For example, if every panel (or grouping of studies) is such that within each panel all the effect-size estimates come from studies that use the same CV method, it would not be possible to estimate the influence on the effect-size of different CV methods due to perfect correlation between the dummy for the CV method and the FE dummy for the panel.

(c) The FE-model estimator employs a fixed intercept for each group. To do this in a reliable way, each group must contain several observations. Our survey of 130 meta-analyses reveals that the median number of observations per primary study is only three. Further, if each primary study is to be used as the grouping criterion, it is likely that some groups will have only one effect-size estimate. The FE software will eliminate these observations from the estimation procedure, in exactly the same way in which an observation-specific dummy eliminates an observation from a traditional regression (Kennedy, 2008, p. 236). Consequently, to employ the FE estimation procedure in meta-analysis, the meta-analyst may need to group the observations such that each grouping contains at least two observations, preferably many more. In some applications this is not realistic, prompting analysts to create groupings so artificial that net benefits from employing this estimation procedure are questionable (see Aigner 1973 for a related discussion).[23] Clearly, some tradeoffs are involved with the use of the FE panel model.

We have three recommendations from this discussion of correlated effect-size estimates. First, if OLS is to be used, robust standard errors must be employed for inference.[24] Second, the advantages of random-effects estimation are so strong that this estimation procedure should be employed unless a very strong case can be made for its inappropriateness. There is no good reason to prefer any one of the hierarchical/multilevel, panel, or clustering software; use whichever is most convenient. Third, although for meta-analysts the most common way of creating levels, panels, or clusters is to group effect-size estimates on the basis of the primary study that created them, in many instances the grouping criterion is not obvious. In the spirit of sensitivity analysis, as illustrated by Rosenberger and Loomis (2000b), empirical results need to be reported for several relevant stratifications of the data.

---

[23] Clustering software is another way of producing the estimates presented above for hierarchical/multilevel and panel data. Cameron and Trivedi (2005) have a good exposition of clustering as an estimation issue, including its similarities with hierarchical/multilevel and panel-data models. They identify Stata and Sudaan as software particularly strong for cluster estimation. An advantage of the clustering software is that it more readily produces robust standard error estimates if OLS is employed.

[24] To be clear here, this does not mean using the White heteroskedasticity-consistent variance-covariance estimator with OLS in econometric software. What is needed is a robust variance-covariance matrix calculation that incorporates the specific nature of the correlation in meta-analyses, usually referred to as cluster-robust standard errors; see Cameron and Trivedi (2005, p.834). Wooldridge (2003) surveys this problem, noting difficulties when the number of clusters is small.

### 2.4. Outlier Analysis

It is evident in most meta-analyses that the range of estimates from primary studies is very large, so much so that one cannot help but wonder if some estimates are unrepresentative or overly influential. This is motivation for discarding of outlying observations or for use of a robust estimation procedure such as OLS with Huber weights (e.g., Viscusi and Aldy 2003).[25] Graphical presentation of the data is important as support for any treatment of outliers as well as formal analysis of residuals (e.g., Espey 1996; Johnson et al. 1997). Our recommendation on this issue is a comparison of results from a robust estimator to those of a first-choice estimator such as weighted least-squares: if the results differ by an uncomfortable amount, re-think the analysis. This advice simply follows the tenth commandment of applied econometrics (Kennedy 2002, 2008): conduct a sensitivity analysis. This recommendation extends to other issues: are the results of the meta-analysis affected substantially by inclusion of multiple observations from each primary study, by estimation using more homogeneous subsamples or exclusion of outliers, by adjusting for heteroskedasticity, by using a different functional form or weighting scheme, or by allowing for random effects? With the above results as a foundation, we next turn to an assessment of 130 meta-analyses in environmental and resource economics.

## 3. A SUMMARY OF 130 META-ANALYSES

Although meta-analysis has seen widespread use, this has not been without criticism (Berk 2007; Elvik 2005; Goldfarb and Stekler 2002; Smith and Pattanayak 2002). Implicit in any meta-analysis is the assumption that the primary studies are similar enough that they can be usefully combined or analyzed. Common criticisms are that meta-analysis compares "apples to oranges," pools primary studies of varying quality, engages in double counting, and suffers from publication bias. These objections are dealt with using appropriate samples and techniques, but require that the investigator prepare an analysis that addresses a variety of problems. In order to outline various methods, this section summarizes 130 meta-analyses. A more detailed assessment of 19 selected studies is contained in the next section.

### 3.1 Selection and Coding of 130 Meta-Analyses

We conducted a search of the relevant literature in environmental and resource economics using a variety of printed and electronic resources, including EconLit, Agricola, AgEconSearch, JSTOR, RePEc, SSRN, Ingenta, ProQuest, Questia, Social Sciences Citation Index, WebFerret, numerous Google searches, ancestry searches, on-line valuation databases (ARIES, EVCBN, EVRI, Envalue, GEVAD,

---

[25] Huber-weighting minimizes a weighted sum of squared errors with weights decreasing as the error become larger. Other formal techniques that have been employed for outliers are trimmed least-squares (Desvousges et al. 1998; Johnson et al. 1997; Murphy et al. 2005) and the minimum absolute deviation estimator (Smith and Huang 1995).

RED, TRIS), and searches of the web sites of major publishers of academic journals, including Blackwell, Elsevier, Kluwer, Sage, Springer, Taylor & Francis, and Wiley. The "grey literature" was searched using the National Technical Information Service, Dissertation Abstracts, web sites of major government agencies (EPA, Forest Service), Science.gov, and web sites of key academic institutions and authors (e.g., Tinbergen Institute). In all, the literature search process took about three months (October-December 2007). Over 300 studies (list available upon request) were examined, and 115 of these were selected for inclusion in the present paper based on the study's content and methods. Studies were excluded if the meta-analysis was a duplicate, an application of an existing analysis, based solely on a tabular meta-summary, or if the study used rough set analysis and did not perform a meta-regression. We also excluded analyses based on experimental studies. A few studies in our sample provide more than one analysis, so the final sample contains 130 meta-analyses derived from 115 studies. We grouped the analyses into seventeen exclusive areas according to the major topic of the primary studies. The topics and number of analyses in each area are: air pollution (5 meta-analyses); stated preference methodology (12); discount rates (3); endangered species and biodiversity (5); energy markets (12); environmental regulation and economic growth (8); global warming and sustainability (14); hazardous waste and pesticides (7); recreation values for aquatic resources (6); recreation values for forestry resources (4); recreation values for multiple-use resources (8); transportation externalities (9); value of a statistical life (11); value of travel time (6); water demand and supply (7); water quality management (7); and wetlands resources (6).

### 3.2 Quantitative Summary of 130 Meta-Analyses

The selected studies are listed in an appendix (available on-line), while Tables 1 and 2 display basic data on study characteristics. The sample includes 89 peer-reviewed journal articles, 7 book chapters, 20 working papers, and 14 other publications (9 dissertations, 5 government reports).[26] The median analysis was published or prepared in the year 2003. For each of the analyses, we coded over 75 study descriptors for effect-size, primary-data type, basic models, regression methods and results, statistical tests, reporting procedures, and applications. In order to remove inconsistencies, the data were coded twice. As a rough quantitative measure of overall quality or completeness of each analysis, a sum of 55 attributes was calculated (e.g., models, graphs, regressions, applications). Counts of study characteristics are provided in Table 1, while Table 2 provides summary statistics for 10 items, including the completeness score. In most categories, the count data in Table 1 sum to more than 130, reflecting multiple entries for many studies. A typical analysis might include a statement of the meta-regression

[26] The 89 articles appear in 49 different academic journals. The journals that have published four or more meta-analyses are Ecological Economics, Energy Economics, Environmental & Resource Economics, Journal of Policy Analysis and Management, Land Economics, and Water Resources Research.

model, empirical results for two different regression models, one treatment or test for heteroskedasticity, one treatment for non-independence, a graph or analysis of outliers, a list of primary studies, a reporting of some primary data, and one application such as a within-sample prediction. Hence, the average score for completeness in Table 2 has a median value of 9. However, the range is from a minimum of 2.0 to a maximum score of 23, suggesting that there is considerable room for improvement in sampling methods, estimation, and reporting procedures in many analyses.

**1. Sample selection and reporting procedures**. Table 1 shows that most analyses include a list of the primary studies (114 of 130 analyses), but fewer provide a statement of selection criteria (54), and many of these are very terse. Twenty-six analyses provide all of the primary data and 48 provide some data. Publication of at least some data is a recommended procedure, such as number of primary estimates per study, effect-size estimates, and context variables. For effect-sizes, stated preference data are examined in 29 analyses, while combinations of stated and revealed preference data are used in 26 analyses. For primary data, survey-based data from stated preference (SP) and travel cost studies are most common (46 of 130 analyses), followed by combinations of data types (39). The use of combination data and aggregate data (16 analyses) raise issues of heterogeneity and non-independence, both within and between studies (see Smith and Pattanayak 2002).

**2. Basic data summary**. Seventy-one studies provide an explicit statement of the meta-regression model and 35 analyses provide an explicit statement of the general model that generates the data, such as the hedonic price model for housing or the regression model used in environmental Kuznets Curve studies. Forty-seven analyses provide plots of the data. These practices aid the selection of explanatory variables and detection of outliers. However, only 14 studies estimate weighted means. As discussed above, we recommend more extensive examinations of the basic data prior to estimation of meta-regressions. Several recent economic studies of publication bias provide good examples of descriptive analyses coupled with meta-regressions (see Roberts and Stanley 2005).

**3. Primary data heterogeneity**. Table 1 shows that 31 studies estimate only a simple OLS model, which is clearly inadequate. Another 35 studies estimate an OLS model with White or Huber-White standard errors for heteroskedasticity. Weighted least-squares are used in 34 analyses, but the weights vary. Five analyses report inverse standard errors as weights and ten use the variances. Thirteen analyses use weights based on primary study sample sizes, while some studies experiment with different weights. Fixed-effects panel models are estimated by 12 analyses, while 17 analyses estimate random-effects panel models (overlap exists here). Beyond these methods, six analyses use a multilevel model, 10 studies use maximum likelihood estimation such as the mixed-effect-size model, and 12 studies use methods for limited and qualitative dependent variables such as an analysis of coefficient significance levels via probit regressions (e.g., Jeppesen et al. 2002; Kiel and Williams 2007; Smith and Huang 1993).

The “other category” includes Box-Cox models, various robust methods for outliers, and two studies that use Bayesian methods (Kochi et al. 2006; Moeltner et al. 2007).

Table 2 shows that the average analysis includes 42 primary studies and 182 primary observations, which is a mean value of 6.5 observations per study (median of 3). In the average regression, there are 14 independent variables, but only half of these are statistically significant (5% confidence level or better). The mean value of the adjusted R-square is 0.48 (median of 0.45). The degree of fit of the average regression suggests the existence of unexplained heterogeneity, which might be reduced by greater use of panel-data methods or meta-regressions on more homogeneous subsamples.

**4. Treatment of heteroskedasticity**. Many meta-analyses provide some treatment of heteroskedasticity, but Table 1 shows that 43 of 130 studies do not report any treatment. The most common treatments are weighted regressions (42 analyses) and use of White (25) or Huber-White (20) standard errors. However, seven studies use the Newey-West estimator, which is designed for stationary time-series data. Use of this estimator in the context of a meta-regression is not appropriate. Other types of treatment are reported in 20 studies, such as estimation using more homogeneous subsamples or bootstrap estimation of coefficient standard errors (e.g., Smith and Huang 1995).

**5. Lack of independence, outliers, and applications**. Forty-seven of 130 studies report no treatment for data dependencies or correlation. The most common treatments are use of a single observation per primary study (27 studies), Huber-White standard errors (20 studies), and panel-data methods (28 studies). Other “treatments” of non-independence include omission of duplicate studies and calculation of intra-study error correlations (e.g., Espey 1996). Outliers are examined in 37 of 130 studies, but only 15 studies report a residual analysis. Among the additional statistical tests, the most common is a model specification test (e.g., Breusch-Pagan Lagrange Multiplier test for panel effects, Hausman test), but 74 studies report no additional tests beyond basic summary statistics for the regressions. We recommend greater use of formal tests for heteroskedasticity, independence, specification, and outliers. Fifty studies include a variable that captures other “quality” aspects of the primary studies, such as a binary variable for unpublished papers. Finally, a within-sample prediction is the most common application of the empirical results (42 studies).

**6. Summary**. Tables 1 and 2 demonstrate substantial diversity in methods and procedures used for meta-analyses in the area of environmental and resource economics. In part, this diversity reflects the relative newness of meta-regression analysis as part of the economist’s tool-kit of econometric methods. However, we share a concern with the EPA Work Group (2006) that the methods used by economists do not fully account for a variety of data problems. As expressed by the Work Group (EPA 2006, p. 6), “important issues concern analyses of subgroups of studies, weighting of individual study results, and the use of regression in meta-analysis.” These concerns can be summarized by noting that 43 of 130 studies ignore possible heteroskedasticity, 47 studies ignore possible correlated observations, and 21 of these studies ignore both problems. In order to further assess these concerns as well as document possible solutions, we next turn to a more detailed assessment of 19 meta-analyses.

## 4. AN ASSESSMENT OF 19 META-ANALYSES

Nineteen meta-analyses were selected for a more detailed assessment. We used the three VSL studies examined by the EPA Work Group and one study from each of the other sixteen topical categories. In general, we have selected meta-analyses that illustrate more advanced econometric methods and explicit consideration of one or more of the major issues discussed above. This selection was aided by the completeness score, but we considered only peer-reviewed articles and book chapters for inclusion in this section. Our discussion of issues is organized following the assessment of the EPA Work Group. Table 3 displays selected aspects of the 19 meta-analyses (a longer summary is available). We have selected eight items for discussion and assessment.

### 4.1 Effect-Size and Data Type

Like any research project, a meta-analysis should begin with a statement of the topic to be investigated or the question to be addressed (EPA 2006, p. 20). This goal determines the population of primary studies and the relevant effect-size. In general, the stated goal also determines criteria for inclusion and exclusion of primary studies and might determine subpopulations of interest, such as by geographic location or valuation method. Next, with the objectives specified, the analyst must identify the primary studies to include in the meta-analysis. One characteristic of a complete analysis is an explicit statement of criteria for relevant studies and a predetermined procedure or protocol for retrieval of the primary studies (EPA 2006, p. 10). This should include criteria for inclusion and exclusion of studies (EPA 2006, p. 18). While many meta-analyses in environmental economics provide some guidance in this area, most fall well short of an adequate statement. It is simply difficult to judge the thoroughness of the analysis without detailed information about the process of primary study retrieval. References to literature reviews, bibliographies, relevant journals, electronic databases and search engines are important aspects of a meta-analysis, but few analyses in economics make a serious effort to provide sufficient details on the retrieval and coding process. Only a few of the 19 studies in Table 3 provide clear criteria and procedures for the data collection phase, especially Asenso-Boadi et al. (2008), Johnston et al. (2006), and Van Houtven et al. (2007). A list of excluded studies is reported in Nelson (2004).

A second feature of a complete analysis is a detailed list of the primary studies contained in the sample and a reporting of some or all of the primary data, which facilitates replication (EPA 2006, p. 9). Complete references should be provided, if necessary using an on-line format. Only six of the 19 studies contain all of the data or provide it on-line. Many of the studies use various combinations of primary data, e.g., contingent valuation surveys and hedonic wage studies. The Work Group report was critical of VSL studies for failure to provide separate analyses for these two types of estimates (EPA 2006, p. 22), but this criticism applies only to Kochi et al. (2006).[27] As emphasized by Smith and Pattanayak (2002), meta-analysis studies also need to adopt a higher standard of consistency in the economic concepts being studied. Several of the 19 analyses can be faulted using this criterion. For example, travel cost (TC) studies use Marshallian surplus for benefit valuation, while CV studies use a Hicksian benefit measure. Four of the 19 studies use estimates from both TC and CV surveys, including Bateman and Jones (2003), Brander et al. (2006), Johnston et al. (2006), and Rosenberger and Loomis (2000a, 200b). We recommend analysis of consistent effect-size estimates. A similar standard should be adopted for the types of environmental commodities considered (Smith and Pattanayak 2002).

### 4.2 Number of Primary Studies and Number of Observations

The Work Group expressed concern that VSL analyses used "multiple VSL estimates from the same study, and therefore from the same data set, [which] should be combined cautiously or not at all" (EPA 2006, p. 26). The 19 studies in Table 3 reveal a variety of approaches for this issue. Four studies use a single observation per study or construct relatively independent data sets, including the VSL study by Viscusi and Aldy (2003). The methods used by Kochi et al. (2006) are insightful as they use 197 primary observations to construct 60 homogeneous groups. In general, the other 15 studies illustrate a major concern with meta-analyses in economics, which is the use of multiple observations from each primary study. The mean number of observations per primary study is 7.5 (median is 3.5), so the 19 analyses are representative of our larger sample. It is not uncommon for a few primary studies to provide the bulk of the observations. For example, in Rosenberger and Loomis (2000a), one primary study of recreation values provides 134 observations. We recommend that additional details are reported for each primary study that provides numerous observations. Because many analysts delete observations selectively, it is usually impossible to tell how many studies are represented in the regressions. Better reporting of the data employed at each stage of the analysis is necessary (e.g., van Kooten et al. 2004). Further, the Work Group also expressed concern that the primary studies employ common data, *viz*., VSL primary studies "draw on a limited number of data sets . . . as sources of the wage data . . . to the extent that they overlap, the VSL estimates will not provide independent information" (EPA 2006, p. 15). While

[27] A possible "quality" indicator is publication status of each primary study, included in several studies as an explanatory variable (e.g., Van Houtven et al. 2007; van Kooten et al. 2004, 2007). A number of studies also include a variable for the year of publication as a control for improvements in stated preference methods. Some analyses included in Table 1 restrict the sample to only primary studies published in few major journals or rely exclusively on one literature survey for study retrieval. These limited search procedures are not recommended (EPA 2006, p. 18).

this concern does not apply to all areas of analysis, it is important that future meta-analyses determine if primary studies use overlapping samples. This is likely to be especially important where the basic data are aggregate time-series or utilize public surveys (e.g., Current Population Survey, Panel Study of Income Dynamics). Controlling for non-independence may require a more detailed examination of the primary data that goes beyond just including an explanatory variable for the relevant survey (EPA 2006, p. 16).

### 4.3 Preliminary Meta-Analysis

Most studies provide summary statistics and a list of the explanatory variables as part of the data description. However, key preliminary steps often are omitted. The Work Group recommended calculation of weighted means and graphical plots of the data and residuals (EPA 2006, pp. 27-8). Weighted means are reported in only two of 19 studies. Only six studies provide graphical displays of the primary data, ranging from simple scatter diagrams to non-linear loess plots. Graphs of effect-size distributions often reveal highly skewed data (e.g., Espey and Espey 2004; Rosenberger and Loomis 2000a). Graphs are useful in order to detect outliers or motivate basic relationships as in Brander et al. (2006), Dalhuisen et al. (2003), and Johnson et al. (1997). A brief statement of the primary-study model is helpful to motivate the selection of explanatory variables and achieve the desired consistency of effect-sizes (e.g., Johnson et al. 1997; Mrozek and Taylor 2002; Nelson 2004; Viscusi and Aldy 2003). An explicit statement of the meta-regression model is useful as a summary of the preliminary stage. As noted by the Work Group (EPA 2006, pp. 19-20), it is important to make clear what population is being represented in deriving a combined effect-size and "whether in fact a single value should indeed be the goal." Only a few of the 19 studies are clear about the basic difference between a meta-analysis based on a "fixed-effect-size" model versus a "random-effect-size" model. We recommend that the preliminary stage is the appropriate point to make this determination clear, with supporting graphs and summary statistics that correspond to subpopulations of possible interest.

### 4.4 Estimation of Meta-Regressions

The 19 studies summarized in Table 3 present the full range of meta-regressions estimated by economists, including simple OLS, weighted least-squares, panel models for fixed- and random-effects, and multilevel models. We have emphasized above the necessity to consider the data-generating process and tailor the regression model to the available data. If the analysis is based on a single observation per primary study or uses otherwise independent data, the analysis should include a generalized least-squares model that employs weighted observations (inverse variances, sample sizes). The Work Group emphasized the use of weighted least-squares, but also recommended the use of one observation from each primary study or each independent data set (EPA 2006, p. 26). Due to small samples of studies, it is likely that this approach is not workable for many meta-analyses conducted in economics. Several of the 19 studies use more advanced regression models, including FE and RE panel-data models that account for unobserved heterogeneity and non-independence (e.g., Florax et al. 2005; Johnson et al. 1997; Rosenberger and Loomis 2000b; van Kooten et al. 2004). As discussed above, RE-models should be supplemented with a Hausman specification tests for fixed-effects (e.g., Florax et al. 2005; Jeppesen et al. 2002; Johnson et al. 1997; Johnston et al. 2006; Rosenberger and Loomis 2000b). If the independence assumption seems justified, more complex multilevel models also can be considered (e.g., Bateman and Jones 2003; Johnson et al. 1997). We recommend that future meta-analyses consider more than one specification of the meta-regression model, which needs to be supported by diagnostic tests.

The Work Group also expressed a concern about the use of effect-sizes based on primary study regressions that "do not include the same independent variables" (EPA 2006, p. 10). The issue here is that specifications differ within or across primary studies, so the resulting effect-size estimates are likely biased to different degrees. There are four possible ways of addressing this problem. First, a dummy variable regressor representing each primary-study specification could be included in the meta-regression. However, due to the plethora of specifications in economic studies, this may not be realistic because there are not enough observations with identical specifications to allow estimation to proceed. Second, meta-analysts could argue that explanatory variables omitted from some primary studies are orthogonal to the focal explanatory variable (i.e., the effect-size variable), and so do not bias the estimation. The Work Group expressed this position by noting that knowledge of covariances among the parameter estimates would be less critical if the explanatory variables used in the primary studies are not strongly interrelated (EPA 2006, p. 15). Third, meta-analysts could argue that range of primary specifications is such that some create upward-biased estimates and some create downward-biased estimates, on average producing unbiased effect-size estimates. A fourth option, sample size permitting, is to perform the meta-analysis on homogeneous subsamples according to the primary model specification. Overall, this is an awkward issue for meta-analysis in economics for which no obvious best solution is available.[28]

### 4.5 Summary of Meta-Regression Models

A survey by Florax et al. (2002, p. 8) provides the following general statement of the meta-regression model

$$Y = f(P, X, R, T, L) + e \tag{7}$$

where $Y$ is the effect-size (e.g., hedonic price for noise levels), $P$ is a set of causes of the outcome (aircraft noise levels), $X$ are characteristics of the set of objects affected by $P$ in order to determine the outcome

[28] Monte Carlo studies by Koetse et al. (2005, 2007a) suggest that including dummy variables for primary model specification can correct for omitted variable bias, but not for erroneous effect-size measures.

(age of houses), $R$ is a set of characteristics of the research methods in each primary study (econometric methods), $T$ indicates the time period of the primary study, $L$ is the location of the objects, and $e$ is the error term. This is a typical grouping of explanatory variables in a meta-regression. Analysis of the 19 studies shows that the mean study includes 18 explanatory variables and half of the explanatory variables are statistically significant (5% level). The mean adjusted R-square is 0.460. In part, the relatively small adjusted R-squares reflect the heterogeneity issues discussed above, which are not necessarily solved by including more regressors (EPA 2006, p. 24). The analyst may instead want to focus on homogeneous subsamples that permit a parsimonious model specification. Studies that illustrate this approach include Mrozek and Taylor (2002), Nelson (2004), Rosenberger and Loomis (2000a), and Smith and Osborne (1996). Detecting and treating outliers also is important. Viscusi and Aldy (2003) illustrate the use of Huber weights as a formal treatment of outliers, while several other studies delete outlying observations (e.g., Johnson et al. 1997; Espey and Espey 2004).

As a simple test of model specification and robustness, we used the sample of 130 meta-analyses to calculate correlations between the number of primary studies and the adjusted R-square ($r = -0.280$) and the number of observations per study and the adjusted R-square ($r = 0.226$). The negative correlation shows that larger samples are more diverse, which supports our suggestion that the analysis can be sharpened by using more homogeneous subsamples. However, this increased focus must be traded-off with concerns with publication bias. As emphasized by Moeltner et al. (2007), a similar dilemma exits for the number of regressors. The positive correlation with the number of observations per study also is a cause for concern. It suggests that R-square values can be artificially inflated by the inclusion of similar observations from the same study. As noted by Rosenberger and Loomis (2000b) reducing the complexity of the data facilitates identification of panel stratification, and “meta-analyses may be most capable of providing definitive conclusions when these results are from fairly homogeneous sources” (p. 466).

### 4.6 Summary of Econometrics

Our assessment has produced evidence that a complete meta-analysis must address issues of data heterogeneity, heteroskedasticity, and non-independent or correlated observations. The heterogeneity may be observable or not. Heteroskedasticity is likely to be a concern regardless of the design of the meta-analysis, and needs to be treated explicitly. As discussed above, correlated observations can arise for a number of reasons, including common data sources and multiple observations from each primary study. The 19 studies in Table 3 are among the more complete in the economics literature, but each study can be faulted in one or more ways. Some studies fail to weight the observations or do not report robust estimates of the standard errors (e.g., Zamparini and Reggiani 2007). Some studies do not report panel regression models where non-independence is likely (e.g., Dalhuisen et al. 2003) or fail to consider the possibility of clustered data. Most studies do not include a sufficient number of regression diagnostics (EPA 2006, p. 24) and fail to address the issue of publication bias. We recommend that future meta-analyses learn from these transgressions and present more complete analyses that engage in more advanced econometric testing and sensitivity analysis of empirical results.

### 4.7 Summary of Applications

Developing an understanding of the sources of heterogeneity is the motivation in virtually all meta-analyses in economics. Many meta-analyses go beyond this basic goal and develop one or more applications of the empirical results. Some studies compare their results with earlier meta-analyses on the same topic (e.g., Brander et al. 2006; Dalhuisen et al. 2003; Kochi et al. 2006; Mrozek and Taylor 2002; Nelson 2004; Viscusi and Aldy 2003). A few studies develop tests of publication bias, although these are relatively simple (e.g., Van Houtven et al. 2007; van Kooten et al. 2004). A few studies provide explicit tests of fundamental hypotheses, such as the "scope test" in Smith and Osborne's (1996) study of WTP for visibility at national parks. More advanced applications involve using the meta-regression to produce predicted values, either within-sample predicted values for the dependent variable or out-of-sample forecasts of effect-sizes. The latter application also is known as a "benefit transfer" because the estimated meta-regression is used to obtain forecast values of "benefits" for another location or commodity. An example illustrates both these predictive applications. Van Houtven et al. (2007) perform a meta-analysis for WTP for water quality using 18 primary studies and 131 observations. Most of the primary studies use CV methods for WTP valuation. Van Houtven et al. use the empirical results from a weighted least-squares regression to predict WTP for selected scenarios involving three changes in freshwater quality levels for recreation users (boaters, fishermen, swimmers) and nonusers. Interestingly, the predicted WTPs for users are greater than nonusers by a factor of three. Van Houtven et al. (2007, p. 223) also perform a simple benefit-transfer exercise by comparing their WTP estimates to earlier valuation estimates for national water quality changes. They obtain similar values for the fishable-to-swimmable quality increment, but report that the meta-regression estimates are lower by as much as 50 percent for the boatable-to-fishable quality increment. We recommend that future meta-analyses incorporate similar applications of the results, with a focus on hypothesis tests and comparative results.

### 4.8 Final Comments on the 19 Meta-Analyses

The Work Group found that three VSL meta-analyses suffered from a high degree of heterogeneity in the VSL estimates, dependencies from inclusion of multiple estimates derived from the same underlying data, heteroskedasticity, and insufficient reporting of methods and results. Our survey of 19 meta-analyses demonstrates that these problems are not uncommon, but can be addressed using a variety of econometric techniques available in the literature. First, heterogeneity can be addressed by using more homogeneous samples, careful selection of consistent effect-sizes and study descriptors, panel-data and multilevel methods, and by application of robust methods such as the Huber-White variance estimator. Second, dependencies can be addressed by using panel methods, robust variance estimators, and by sensitivity analysis of alternative samples and estimation methods. Third, heteroskedasticity can be treated by using weighted least-squares and robust variance estimators. Finally, virtually all studies could be improved by better reporting of methods and results, including greater use regression diagnostics and graphical illustrations of primary data and regression residuals.

## 5. BEST-PRACTICE GUIDELINES FOR FUTURE META-ANLYSES

It has been reported that it is easy to do a meta-analysis, but it is difficult to do a good one. We concur, and have emphasized above that the choice of econometric techniques depends on the data at hand. The choice of technique also may reflect the objectives of the meta-analysis, including the intended applications such as a benefit transfer. With these general thoughts in mind, we offer ten sets of guidelines for future meta-analyses in environmental and resource economics:

**1. Problem definition.** Provide a clear statement of the problem or hypothesis to be tested, including a precise definition of the effect-size to be investigated. Discuss the correspondence between the theoretical definition of the effect-size and its operational or empirical definition, with possible assistance from the theoretical models in the primary studies. Ensure that the effect-size measures from the primary studies are all measuring the same thing. Consider and report the context in which the primary studies are conducted that determines how the primary data are generated. Determine if the primary studies use overlapping data, such as public surveys or aggregate data.

**2. Search protocol.** Provide a statement of the search strategy or protocol, including sources used for gathering of published and unpublished primary studies. Construct and report criteria for distinguishing relevant from irrelevant studies, and define *a priori* subpopulations of interest. Describe changes in the sample size as studies or observations are deleted, and clearly report the number of studies and observations in the final sample. If data are missing for some studies, ask the question "why are they missing?" Provide a clear reporting of how missing data were handled.

**3. Coding of data.** Provide sufficient documentation regarding coding, rationale for selection of explanatory variables, and adjustments to effect-sizes. Discuss the consistency of effect-size measures across studies and defend as necessary the deletion of any observations from the primary study sample. Present the categories of variables to be included in the meta-regressions and summarize as necessary the expected signs on these variables. Present a list of the studies included in meta-analysis, with some or all of the primary data (possibly on-line). Indicate the number of observations obtained from each of the primary studies and attempt to acquire the sample size and variance for each effect-size estimate. Assess the possibility that some attributes may vary due solely to a small number of primary studies.

**4. Preliminary meta-analysis.** The preliminary meta-analysis should include the presentation of FES and RES weighted means, summary statistics for explanatory variables, Q-tests for homogeneity, and I-tests for correlatedness of effect-size estimates. Provide a graphical summary and assessment of the effect-size distribution, including identification of outlier observations. Inspect the data for coding errors and other inconsistencies that may be due to outliers. Report and discuss details in any primary study that provides numerous observations relative to other studies.

**5. Data Heterogeneity.** Assess the degree of heterogeneity in the primary-study data and its sources. Indicate and discuss whether the meta-regression analysis will be based on a fixed- or random-effect-size model. Discuss criteria for choosing subsamples and consider samples based on a single observation per primary study as well as multiple observations. For panel-data models, discuss different possible strategies for stratification of the data and use of instrumental variables. Consider estimation of the between-groups regression model as illustrated in Jeppesen et al. (2002).

**6. Meta-regression model.** Choose a meta-regression model that is appropriate for the data and study objectives, such as weighted least-squares and panel-data regressions. Address the issues of data heterogeneity, heteroskedasticity, non-independence of observations, and outliers. Report results from use of robust vs. non-robust methods for standard errors. For each regression, report the number of observations and number of studies represented by these observations. Consider a blend of "testing-down" and "testing-up" strategies for construction of a model specification (Kennedy 2008, pp. 72-6).

**7. Specification tests.** Report results for model specification tests and regression diagnostics, including tests for omitted variables, functional form, outliers, nonspherical errors, and cross-sectional vs. panel models. As a multicollinearity diagnostic, consider reporting variance inflation factors for full and restricted sets of explanatory variables. Statistics for the within-study residual correlations should be reported. Assess the residual quality for each regression, including graphical displays.

**8. Sensitivity analysis**. Report results for a sensitivity analysis of the final model, including results from use of different estimation methods, subsample estimation, deletion of outliers, homogeneous subset regressions, different functional forms or panel methods, and specification of the explanatory variables. Report the statistical significance and *substantive* significance of the regression coefficients (Ziliak and McClosley 2008). Assess the fragility of the final results. Consider the implications for policy analysis of fragile vs. robust estimates.

**9. Publication bias**. Assess the quality of the final results in light of formal tests for publication selection bias. Consider use of several possible tests for publication bias, including funnel plots and more advanced tests such as the funnel asymmetry test. Consult the literature in this area to stay abreast of recent developments and applications (e.g., Roberts and Stanley 2005; Rosenberger and Johnston 2008; Stanley 2005, 2008).

**10. Applications**. Consider presentation of in-sample predictions, out-of-sample benefit transfers, comparisons with prior meta-analyses, and additional tests of fundamental hypotheses. Determine if the results of the meta-analysis support a benefit-transfer application, and discuss why or why not (Smith and Pattanayak 2002).

## 6. CONCLUSIONS

Amongst a "flood of numbers," the promise of meta-analysis is to provide a life-boat that creates order and gives substance to numerous empirical studies, which often present conflicting or wide-ranging estimates of important economic concepts. In this paper, the main problem we have discussed is that the life-boat can be improperly guided or manned, so as to threaten the success of the enterprise. The recommendations provided in this paper are intended to help guide the life-boat along a proper course. To this end, we have discussed the data-generating process that challenges investigators in this area and described meta-regression models that can be used to deal with data heterogeneity, heteroskedasticity, non-independence, and outliers. We have provided a number of recommendations for choices among models, selection of procedures and methods, and reporting of results. Our survey of 130 meta-analysis studies in environmental and resource economics reveals that numerous studies have abused meta-analysis methodology by failing to deal adequately with these issues and related concerns. Nevertheless, the literature does contain studies that provide excellent guidance with regard to econometric methods and associated tests and details required for a complete meta-analysis. The 19 studies summarized in Table 3 are representative of the better studies in each of the seventeen topical categories, but even in these studies there is ample room for improvement. Improvements may not be easy in all cases and our suggestions and guidelines should not be read as prescriptions. Meta-analysis is an art – it involves a difficult balancing act between problem definition, data collection, modeling, and application.

An earlier review by Smith and Pattanayak (2002) focused on two issues associated with meta-analyses that limit the use of such studies for benefit transfers. First, the environmental resources being evaluated may be different, such as meta-analyses that incorporate diverse recreational activities or use widely-varying recreation sites. Second, the valuation concept being combined or analyzed can be different, such as Marshallian or Hicksian consumer surplus. Smith and Pattanayak (2002) are critical of prior meta-analyses on the ground that such studies may be of little use for benefit transfers because a "synthesis requires the ability to define a common concept to be measured" (p. 274). They argue that a "higher standard" must be applied to meta-analyses if the objective is a benefit transfer and propose a method of "structural meta-analysis." This method involves specification of a preference function as a structural model and use of this function to interpret estimates across primary studies that may use different valuation methods, which requires substantial adjustments to the data. The assessment in the present paper goes beyond these two concerns, and focuses on the technique of meta-analysis as currently applied in environmental and resource economics. We conclude that the incomplete or incorrect methods found in many studies, even if applied to consistently measured effect-sizes, will still restrict the application of meta-analytical results to policy problems.

**References**

Aigner, D.J., Regression with a binary independent variable subject to errors of observation, Journal of Econometrics 1, 1973, 49-60.

Asenso-Boadi, F., T.J. Peters, and J. Coast, Exploring differences in empirical time preference rates for health: An application of meta-regression, Health Economics 17, 2008, 235-48.

Baltagi, B.H., Econometric Analysis of Panel Data, 3rd ed. (New York: Wiley, 2005).

Banzhaf, H.S. and V.K. Smith, Meta-analysis in model implementation: Choice sets and the valuation of air quality improvements, Journal of Applied Econometrics 22, 2007, 1013-31.

Bateman, I.J. and A.P. Jones, Contrasting conventional with multi-level modeling approaches to meta-analysis: Expectation consistency in U.K. woodland recreation values, Land Economics 79, 2003, 235-58.

Bellavance, F., G. Dionne, and M. Lebeau, The value of a statistical life: A meta-analysis with a mixed effects regression model, Working paper 06-12, Ecole des HEC Montreal, 2007.

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

Berk, R., Statistical inference and meta-analysis, Journal of Experimental Criminology 3, 2007, 247-70.

Bickel, R., Multilevel Analysis for Applied Research (New York: Guilford Press, 2007).

Brander, L.M., R.J.G.M. Florax, and J.E. Vermaat, The empirics of wetland valuation: A comprehensive summary and a meta-analysis of the literature, Environmental & Resource Economics 33, 2006, 223-50.

Breusch, T.S., G.E. Mizon, and P. Schmidt, Efficient estimation using panel data, Econometrica 57, 1989, 695-700.

Brons, M., P. Nijkamp, E. Pels, and P. Rietveld, A meta-analysis of the price elasticity of gasoline demand: A SUR approach, Energy Economics (in press, 2007).

Brouwer, R., I.H. Langford, I.J. Bateman, and R.K. Turner, A meta-analysis of wetland contingent valuation studies, Regional Environmental Change 1, 1999, 47-57.

Brown, M. and A. Forsythe, Robust tests for the equality of variances, Journal of the American Statistical Association 69, 1992, 364-67.

Cameron, A.C. and P.K. Trivedi, Microeconometrics: Methods and Applications (Cambridge: Cambridge University Press, 2005).

Cavlovic, T.A., K.H. Baker, R.P. Berrens, and K. Gawande, A meta-analysis of environmental Kuznets curve studies, Agricultural and Resource Economics Review 29, 2000, 32-42.

Christensen, P., Topics in Meta-Analysis: A Literature Survey (Oslo: Institute of Transport Economics, 2003).

Cooper, H. and L.V. Hedges (eds.), The Handbook of Research Synthesis (New York: Russell Sage Foundation, 1994).

Dalhuisen, J.M., R.J.G.M. Florax, H.L.F. de Groot, and P. Nijkamp, Price and income elasticities of residential water demand: A meta-analysis, Land Economics 79, 2003, 292-308.

Daniel, V.E., R.J.G.M. Florax, and P. Rietveld, Flooding risk and housing values: An economic assessment of environmental hazard, Working paper 07-02, Purdue University, 2007.

Day, B., A meta-analysis of wage-risk estimates of the value of statistical life, in European Commission, Benefits Transfer and the Economic Valuation of Environmental Damages (Brussels: EC, 1999), pp. 1-29.

de Blaeij, A., R.J.G.M. Florax, P. Rietveld, and E. Verhoef, The value of statistical life in road safety: A meta-analysis, Accident Analysis and Prevention 35, 2003, 973-86.

Debrezion, G., E. Pels, and P. Rietveld, The impact of railway stations on residential and commercial property value: A meta-analysis, Journal of Real Estate Finance and Economics 35, 2007, 161-80.

Desvousges, W.H., F.R. Johnson, and H.S. Banzhaf, Environmental Policy Analysis with Limited Information: Principles and Applications of the Transfer Method (Cheltenham: Elgar, 1998).

Elvik, R., Can we trust the results of meta-analyses? Transportation Research Record 1909, 2005, 221-29.

Espey, J.A. and M. Espey, Turning on the lights: A meta-analysis of residential electricity demand elasticities, Journal of Agricultural and Applied Economics 36, 2004, 65-81.

Espey, M., Explaining the variation in elasticity estimates of gasoline demand in the United States, Energy Journal 17, 1996, 49-60.

Florax, R.J.G.M., Accounting for dependence among study results in meta-analysis: Methodology and applications to the valuation and use of natural resources, Research memorandum 2002-5, Free University of Amsterdam, 2002a.

Florax, R.J.G.M., Methodological pitfalls in meta-analysis: Publication bias, in R.J.G.M. Florax, P. Nijkamp, and K.G. Willis (eds.), Comparative Environmental Assessment (Cheltenham: Elgar, 2002b), pp. 177-207.

Florax, R.J.G.M., P. Nijkamp, and K.G. Willis (eds.), Comparative Environmental Assessment (Cheltenham: Elgar, 2002).

Florax, R.J.G.M., C.M. Travisi, and P. Nijkamp, A meta-analysis of the willingness to pay for reductions in pesticide risk exposure, European Review of Agricultural Economics 32, 2005, 441-67.

Gelman, A. and J. Hill, Data Analysis Using Regression and Multilevel/Hierarchical Models (Cambridge: Cambridge University Press, 2007).

Ghermandi, A., J.C.J.M. van den Bergh, L.M. Brander, H.L.F. de Groot, and P.A.L.D. Nunes, Exploring diversity: A meta-analysis of wetland conservation and creation, Working paper, Free University of Amsterdam, 2007.

Glass, G.V., Primary, secondary, and meta-analysis of research, The Educational Researcher 10, 1976, 3-8.

Goldfarb, R.S. and H.O. Stekler, Meta-analysis, Journal of Economic Perspectives 16, 2002, 225-26.

Greene, W.H., Econometric Analysis, 6th ed. (Upper Saddle River, NJ: Pearson Prentice Hall, 2008).

Hausman, J.A. and W.E. Taylor, Panel data and unobservable individual effects, Econometrica 49, 1981, 1377-98.

Hedges, L.V., Meta-analysis, Journal of Educational Statistics 17, 1992, 279-96.

Hedges, L.V., Fixed effects models, in H. Cooper and L.V. Hedges (eds.), The Handbook of Research Synthesis (New York: Russell Sage Foundation, 1994), pp. 285-99.

Hedges, L.V. and I. Olkin, Statistical Methods for Meta-Analysis (New York: Academic Press, 1985).

Higgins, J.P.T. and S.G. Thompson, Quantifying heterogeneity in a meta-analysis, Statistics in Medicine 21, 2002, 1539-58.

Horowitz, J.K. and K.E. McConnell, A review of WTA/WTP studies, Journal of Environmental Economics and Management 44, 2002, 426-47.

Hox, J.J. and E.D. de Leeuw, Multilevel models for meta-analysis, in S.P. Reise and N. Duan (eds.), Multilevel Modeling: Methodological Advances and Applications (Mahwah, NJ: Lawrence Erlbaum, 2003), pp. 90-111.

Hunter, J.E. and F.L. Schmidt, Methods of Meta-Analysis: Correcting Error and Bias in Research Findings (Thousand Oaks, CA: Sage, 2004).

Jarrell, S.B. and T.D. Stanley, A meta-analysis of the union wage gap, Industrial and Labor Relations Review 44, 1990, 54-67.

Jeppesen, T., J.A. List, and H. Folmer, Environmental regulations and new plant location decisions: Evidence from a meta-analysis, Journal of Regional Science 42, 2002, 19-49.

Johnson, F.R., E.E. Fries, and H.S. Banzhaf, Valuing morbidity: An integration of the willingness-to-pay and health-status index literatures, Journal of Health Economics 16, 1997, 641-65.

Johnston, R.J., E.Y. Besedin, and R.F. Wardwell, Modeling relationships between use and nonuse value for surface water quality: A meta-analysis, Water Resources Research 39, 2003, 1363-71.

Johnston, R.J., E.Y. Besedin, R. Iovanna, C.J. Miller, R.F. Wardwell, and M.H. Ranson, Systematic variation in willingness to pay for aquatic resource improvements and implications for benefit transfer: A meta-analysis, Canadian Journal of Agricultural Economics 53, 2005, 221-48.

Johnston, R.J., M.H. Ranson, E.Y. Besedin, and E.C. Helm, What determines willingness to pay per fish? A meta-analysis of recreational fishing values, Marine Resource Economics 21, 2006, 1-32.

Kennedy, P.E., Estimation with correctly interpreted dummy variables in semilogarithmic equations, American Economic Review 71, 1981, 802.

Kennedy, P.E., Sinning in the basement: What are the rules? The ten commandments of applied econometrics, Journal of Economic Surveys 16, 2002, 569-89.

Kennedy, P.E, A Guide to Econometrics, 6th ed. (Malden, MA: Blackwell, 2008).

Kiel, K.A. and M. Williams, The impact of Superfund sites on local property values: Are all sites the same? Journal of Urban Economics 61, 2007, 170-92.

Kochi, I., B. Hubbell, and R. Kramer, An empirical Bayes approach to combining and comparing estimates of the value of a statistical life for environmental policy analysis, Environmental & Resource Economics 34, 2006, 385-406.

Koetse, M.J., R.J.G.M. Florax, and H.L.F. de Groot, Correcting for primary study misspecification in meta-analysis, Tinbergen Institute Discussion Paper TI 2005-029/3, Free University of Amsterdam, 2005.

Koetse, M.J., R.J.G.M. Florax, H.L.F. de Groot, The impact of effect size heterogeneity on meta-analysis: A Monte Carlo experiment, Tinbergen Institute Discussion Paper TI 2007-052/3, Free University of Amsterdam, 2007a.

Koetse, M.J., H.L.F. de Groot, and R.J.G.M. Florax, Capital-energy substitution and shifts in factor demand: A meta-analysis, Energy Economics (in press, 2007b).

Lindhjem, H. and S. Navrud, How reliable is meta-analysis for international benefit transfers? Ecological Economics 66, 2008, 425-35.

Lipsey, M.W. and D.B. Wilson, Practical Meta-Analysis (Thousands Oaks, Ca: Sage, 2001).

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

Messer, K.D., W.D. Schultze, K.F. Hackett, T.A. Cameron, and G.H. McClelland, Can stigma explain large property value losses? Psychology and economics of Superfund, Environmental & Resource Economics 33, 2006, 299-324.

Moeltner, K., K.J. Boyle, and R.W. Paterson, Meta-analysis and benefit transfer for resource valuation–Addressing classical challenges with Bayesian modeling, Journal of Environmental Economics and Management 53, 2007, 250-69.

Moulton, B. R., An illustration of a pitfall in estimating the effect of aggregate variables on microeconomic units, Review of Economics and Statistics 72, 1990, 334-38.

Mrozek, J.R. and L.O. Taylor, What determines the value of life? A meta-analysis, Journal of Policy Analysis and Management 21, 2002, 253-70.

Mulatu, A., R.J.G.M. Florax, and C.A. Withagen, Environmental regulation and competitiveness: An exploratory meta-analysis, in C. Böhringer and A. Löschel (eds.), Empirical Modeling of the Economy and the Environment (Berlin: Physica Verlag, 2003), pp. 23–54.

Murphy, J.J., P.G. Allen, T.H. Stevens, and D. Weatherhead, A meta-analysis of hypothetical bias in stated preference valuation, Environmental & Resource Economics 30, 2005, 313-25.

Nelson, J.P., Meta-analysis of airport noise and hedonic property values: Problems and prospects, Journal of Transport Economics and Policy 38, 2004, 1-28.

Oltmer, K., P. Nijkamp, R. Florax, and F. Brouwer, A meta-analysis of environmental impacts of agri-environmental policies in the European Union, Tinbergen Institute Discussion Paper TI 2000-058/3, 2000.

Patuelli, R., P. Nijkamp, and E. Pels, Environmental tax reform and the double dividend: A meta-analytical performance assessment, Ecological Economics 55, 2005, 564-83.

Poe, G.L. K.J. Boyle, and J.C. Bergstrom, A preliminary meta analysis of contingent values for ground water quality revisited, in J.C. Bergstrom, K.J. Boyle, and G.L. Poe (eds.), The Economic Value of Water Quality (Cheltenham: Elgar, 2001), pp. 137-62.

Quantitative Micro Software, EViews 6.0 User's Guide, Part II (Irvine, CA: QMS, 2007).

Rabe-Hesketh, S. and A. Skrondal, Multilevel and Longitudinal Modeling Using Stata, 2nd ed. (College Station, TX: Stata Press, 2008).

Raudenbush, S.W., Random effects models, in H. Cooper and L.V. Hedges (eds.), The Handbook of Research Synthesis (New York: Russell Sage Foundation, 1994), pp. 301-21.

Raudenbush, S.W. and A.S. Bryk, Hierarchical Linear Models: Applications and Data Analysis Methods, 2nd ed. (Thousand Oaks, CA: Sage, 2002).

Ringquist, E.J., Assessing evidence of environmental inequities: A meta-analysis, Journal of Policy Analysis and Management 24, 2005, 223-47.

Roberts, C.J. and T.D. Stanley, Meta-Regression Analysis: Issues of Publication Bias in Economics (Oxford: Blackwell Publishing, 2005).

Rosenberger, R.S. and J.B. Loomis, Using meta-analysis for benefit transfer: In-sample convergent validity tests of an outdoor recreation database, Water Resource Research 36, 2000a, 1097-107.

Rosenberger, R.S. and J.B. Loomis, 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, 2000b, 459-70.

Rosenberger, R.S. and R.J. Johnston, Selection effects in meta-analysis and benefit transfer: Avoiding unintended consequences, Working paper, Oregon State University, 2008.

Rosenberger, R.S. and T.D. Stanley, Publication bias in the recreation use value literature: A preliminary investigation, Working paper, Oregon State University, 2007.

Rothstein, H.R., A.J. Sutton, and M. Borenstein, Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments (New York: Wiley, 2005).

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

Saxonhouse, G.R., Estimated parameters as dependent variables, American Economic Review 66, 1976, 178-83.

Schulze, R., Meta-Analysis: A Comparison of Approaches (Cambridge, MA: Hogrefe & Huber, 2004).

Smith, V.K. and J-C. Huang, Hedonic models and air pollution: Twenty-five years and counting, Environmental & Resource Economics 3, 1993, 381-94.

Smith, V.K. and J-C. Huang, Can markets value air quality? A meta-analysis of hedonic property value models, Journal of Political Economy 103, 1995, 209-27.

Smith, V.K. and Y. Kaoru, What have we learned since Hotelling's letter? A meta-analysis, Economics Letters 32, 1990a, 267-72.

Smith, V.K. and Y. Kaoru, Signals or noise? Explaining the variation in recreation benefit estimates, American Journal of Agricultural Economics 72, 1990b, 419-33.

Smith, V.K. and L.L. Osborne, Do contingent valuation estimates pass a 'scope' test? A meta-analysis, Journal of Environmental Economics and Management 31, 1996, 287-301.

Smith, V.K. and S.K. Pattanayak, Is meta-analysis a Noah's Ark for non-market valuation? Environmental & Resource Economics 22, 2002, 271-96.

Stanley, T.D., Wheat from chaff: Meta-analysis as quantitative literature review, Journal of Economic Perspectives 15, 2001, 131-50.

Stanley, T.D., Beyond publication bias, Journal of Economic Surveys 19, 2005, 309-45.

Stanley, T.D., Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection, Oxford Bulletin of Economics and Statistics 70, 2008, 103-27.

Stanley, T.D. and S.B. Jarrell, Meta-regression analysis: A quantitative method of literature surveys, Journal of Economic Surveys 3, 1989, 161-70.

StataCorp, Stata Statistical Software, Release 10, Base Reference Manual, Vol. 3 (College Station, TX: Stata Press, 2007).

Sutton, A.J., K.R. Abrams, D.R. Jones, T.A. Sheldon, and F. Song, Methods for Meta-Analysis in Medical Research (Chichester, UK: Wiley, 2000).

U.S. Environmental Protection Agency, Guidelines for Preparing Economic Analyses, EPA 240-R-00-003 (Washington, DC: EPA, 2000).

U.S. Environmental Protection Agency, Report of the EPA Work Group on VSL Meta-Analysis, Report NCEE-0494, National Center for Environmental Economics, EPA, 2006. Available at http://yosemite.epa.gov/ee/epa/eerm.nsf/vwRepNumLookup/EE-0494?OpenDocument

U.S. Environmental Protection Agency, SAB Advisory on EPA's Issues in Valuing Mortality Risk Reduction, EPA-SAB-08-001, EPA, 2007. Available at http://yosemite.epa.gov/sab/sabpeople.nsf/WebCommittees/BOARD

van den Bergh, J.C.J.M., K.J. Button, P. Nijkamp, G.C. Pepping, Meta-Analysis in Environmental Economics (Dordrecht: Kluwer, 1997).

Van Garderen, K.J. and C. Shah, Exact interpretation of dummy variables in semilogarithmic equations with estimation uncertainty, Econometrics Journal 5, 2002, 149-59.

Van Houtven, G., J. Powers, and S.K. Pattanayak, Valuing water quality improvements in the United States using meta-analysis: Is the glass half-full or half-empty for national policy analysis? Resource and Energy Economics 29, 2007, 206-28.

van Kooten, G.C., A.J. Eagle, J. Manley, and T. Smolak, How costly are carbon offsets? A meta-analysis of carbon forest sinks, Environmental Science & Policy 7, 2004, 239-51.

van Kooten, G.C., S. Laaksonen-Craig, and Y. Wang, Costs of creating carbon offsets via forestry activities: A meta-regression analysis, Working paper 2007-03, University of Victoria, 2007.

Viscusi, W.K. and J.E. Aldy, The value of a statistical life: A critical review of market estimates throughout the world, Journal of Risk and Uncertainty 27, 2003, 5-76.

Walsh, R.G., D.M. Johnson, and J.R. McKean, Issues in nonmarket valuation and policy applications: A retrospective glance, Western Journal of Agricultural Economics 14, 1989, 178-88.

Walsh, R.G., D.M. Johnson, and J.R. McKean, Nonmarket values from two decades of research on recreation demand, in A. Link and V.K. Smith (eds.), Advances in Applied Microeconomics, Vol. 5 (Greenwich, CN: JAI Press, 1990), pp. 167-93.

Waters, W.G. II, Values of travel time savings in road transport project evaluation, in D. Hensher, J. King, and T.H. Oum (eds.), World Transport Research, Vol. 3: Transport Policy (Oxford: Pergamon, 1996), pp. 213-23.

Weitzman, M.L. and D.L. Kruse, Profit sharing and productivity, in A. Blinder (ed.), Paying for Productivity: A Look at the Evidence (Washington, DC: Brookings Institution, 1990), pp. 95-140.

Won, D-H., J.B. Braden, and L.O. Taylor, The economic impact of contaminated and noxious sites: A meta-analysis, Working paper, University of Illinois, 2007.

Wood, J.A., Methodology for dealing with duplicate study effects in a meta-analysis, Organizational Research Methods 11, 2008, 79-95.

Wooldridge, J. M., Cluster-sample methods in applied econometrics, American Economic Review Papers and Proceedings 93, 2003, 133-38.

Zamparini, L. and A. Reggiani, Meta-analysis and the value of travel time savings: A transatlantic perspective in passenger transport, Networks and Spatial Economics 7, 2007, 377-96.

Ziliak, S.T. and D.N. McCloskey, The Cult of Statistical Significance (Ann Arbor: University of Michigan Press, 2008)

**Table 1**. Summary of 130 Meta-Analyses (count data)

| Study Attribute (multiple counts possible) | No. of Studies | Study Attribute (multiple counts possible) | No. of Studies |
|---|---|---|---|
| Publication type: Total no. | 130 | Heteroskedasticity treatments: | |
| Article | 89 | White se | 25 |
| Chapter | 7 | Newey-West se | 7 |
| Working paper | 20 | Huber-White se | 20 |
| Other | 14 | Explicit weights | 42 |
| | | Other controls | 20 |
| Reporting of primary studies: | | No treatment reported | 43 |
| List of studies (incl. on-line) | 114 | | |
| Selection criteria provided | 54 | Explicit weights used: | |
| Some data provided | 48 | Std. errors | 5 |
| All data provided | 26 | Variances | 10 |
| No data provided | 56 | Primary study sample size | 13 |
| | | No. obs. from primary study | 14 |
| Primary effect-size from: Total no. | 130 | Other | 7 |
| Stated preferences | 29 | | |
| Hedonic price/wage | 19 | Non-independence treatments: | |
| Travel cost | 6 | Single ob. per primary study | 27 |
| Combinations of SP, HP, or TC | 26 | Average of obs. | 9 |
| Elasticity | 20 | Huber-White se | 20 |
| Cost-based | 14 | Panel model – fixed effects | 12 |
| Discount rate | 3 | Panel model – random effects | 16 |
| Other | 13 | Multilevel model | 6 |
| | | Other treatment | 20 |
| Primary data type: Total no. | 130 | No treatment reported | 47 |
| Aggregate data & public surveys | 16 | | |
| Model simulations | 9 | Specification tests provided: | |
| Micro-data | 20 | Omitted variables | 9 |
| Surveys | 46 | Model specification | 21 |
| Combinations of data types | 39 | Functional form | 7 |
| | | Outlier tests | 6 |
| Preliminary analysis: | | Nonspherical errors | 10 |
| Primary model stated | 35 | Homogeneity | 17 |
| Meta-analysis model stated | 71 | Other tests | 5 |
| Graphs of data | 47 | None reported | 74 |
| Weighted means | 14 | | |
| ANOVA model | 15 | Additional tests: | |
| | | Outliers omitted or modeled | 37 |
| Regression models estimated: | | Multicollinearity examined | 23 |
| OLS only – no se corrections | 31 | Residuals analyzed | 15 |
| OLS & other regressions | 19 | | |
| OLS w/White or Huber-White se | 35 | Special variables included: | |
| OLS w/Newey-West se | 6 | Sample size variable | 11 |
| Weighted GLS | 34 | Quality variable | 50 |
| Panel model – fixed effects | 12 | | |
| Panel model – random effects | 17 | Applications of analysis results: | |
| Multilevel model | 6 | Out-of sample benefit transfer | 18 |
| Maximum likelihood estimation | 10 | Within-sample prediction | 42 |
| Probit/logit/tobit models | 12 | Publication bias analyzed | 15 |
| Other (e.g., Box-Cox, Huber) | 14 | Comparison w/ other meta-analyses | 17 |
| | | Other comparisons & tests (e.g., scope) | 28 |
| | | Only reports basic results | 50 |

Abbreviations: SP=stated preference; GLS=generalized least squares; HP=hedonic price; OLS=ordinary least-squares; se= standard error; TC=travel cost

**Table 2**. Quantitative Summary of 130 Meta-Analyses

| **Study Attribute** | **Mean (sd)** | **Median** | **Minimum** | **Maximum** |
|---|---|---|---|---|
| Study year | 2002 (4.5) | 2003 | 1989 | 2007 |
| | | | | |
| No. of primary studies incl. | 42 (36) | 33 | 5 | 228 |
| No. of primary obs. incl. | 182 (223) | 92 | 9 | 1167 |
| No. obs per primary study | 6.5 (12) | 3 | 1 | 77 |
| | | | | |
| No. indep. variables incl. | 14 (8.3) | 12 | 1 | 41 |
| No. significant variables (5%) | 7 (5.7) | 5 | 0 | 35 |
| Pct. significant variables | 50 (25) | 50 | 0 | 100 |
| Degrees of freedom | 166 (220) | 73 | 3 | 1126 |
| Adjusted R-sq. | 0.479 (0.238) | 0.448 | 0 | 0.970 |
| | | | | |
| Completeness rating score – sum of 55 attributes in Table 1 | 9.4 (3.9) | 9 | 2 | 23 |

Notes: If unavailable in the meta-analysis study, adjusted R-square values were calculated by the authors. All table entries reflect the selection of one regression as representative of each study's empirical results.

**Table 3**. Narrative Summary of 19 Meta-Analyses

| Meta-Analysis & Topic | Summary of Study |
|---|---|
| Asenso-Boadi et al. (2008); discount rates | **Effect-size:** Societal time preference rates related to health care; combined data types, so context varies. **No. studies & no. of obs.:** 8 studies; 14 random data points. **Meta-regressions estimated:** WLS w/ inverse standard error wts. **Score:** 8 |
| Bateman & Jones (2003); forestry recreation | **Effect-size:** Value per visit for U.K. woodland recreation; CV & TC studies, so valuation varies. **No. studies & no. of obs.:** 30 studies of 21 forests; 77 obs. **Meta-regressions estimated:** OLS for CV estimates & full sample; multilevel model. **Score:** 8 |
| Brander et al. (2006); wetlands | **Effect-size:** Wetland value per hct., but valuation varies. **No. studies & no. of obs.:** 80 studies; 202 obs. **Meta-regressions estimated:** OLS w/ White standard errors. **Score:** 9 |
| Dalhuisen et al. (2003); water demand | **Effect-size:** Price & income elasticity of demand for water. **No. studies & no. of obs.:** 64 studies; 296 price & 161 income elasticities. **Meta-regressions estimated:** OLS; OLS w/ White standard errors; Box-Cox model. **Score:** 14 |
| Espey & Espey (2004); energy markets: electricity | **Effect-size:** Price & income elasticities; combined data types. **No. studies & no. of obs.:** 36 studies; 125 price & 126 income elasticities. **Meta-regressions estimated:** OLS w/ White standard errors; gamma model. **Score:** 9 |
| Florax et al. (2005); pesticide risks | **Effect-size:** Value for pesticide risk reductions, but valuation method and risk change vary. **No. studies & no. of obs.:** 15 studies; 316 obs. **Meta-regressions estimated:** FE and RE panel models w/ sample size wts. **Score:** 14 |
| Jeppesen et al. (2002); environmental regulation & economic growth | **Effect-size:** Environmental regulation effect on firm location; combined data types. **No. studies & no. of obs.:** 11 studies; 368 obs. **Meta-regressions estimated:** OLS; RE panel model; probit model for significance of elasticity. **Score:** 9 |
| Johnson et al. (1997); air pollution health effects | **Effect-size:** WTP to avoid health effects for air pollution; CV surveys. **No. studies & no. of obs.:** 5 studies; 53 obs. **Meta-regressions estimated:** OLS; trimmed least-squares; FE & RE panel models; separate-variances model. **Score:** 20 |
| Johnston et al. (2006); aquatic recreation | **Effect-size:** Value for quality of fishing changes; SP & TC surveys, so valuation method varies. **No. studies & no. of obs.:** 48 studies; 391 obs. **Meta-regressions estimated:** RE panel model w/ Huber-White se & wts. no. obs. for each study; RE no wts. **Score:** 12 |
| Kochi et al. (2006); value of a statistical life | **Effect-size:** VSL; SP & hedonic wage studies, so valuation varies. **No. studies & no. of obs.:** 31 hedonic & 14 CV studies; 197 obs.; 60 homogeneous subsets. **Meta-regressions estimated:** Empirical Bayes model w/ plots of composite distribution. **Score:** 12 |
| Loomis & White (1996); endangered species | **Effect-size:** WTP for 18 endangered species; CV surveys. **No. studies & no. of obs.:** 20 studies w/ 38 indep. obs. for various species. **Meta-regression estimated:** OLS. **Score:** 7 |
| Mrozek & Taylor (2002); value of a statistical life | **Effect-size:** VSL; hedonic wage studies. **No. studies & no. of obs.:** 33 studies; 203 obs. **Meta-regressions estimated:** OLS w/Huber-White standard errors; WLS w/ study obs. as wts. & Huber-White se; WLS w/ t-stat wts.; Box-Cox model. **Score:** 15 |
| Nelson (2004); transport externalities: noise & property values | **Effect-size:** Pct. change in property value per decibel; hedonic studies. **No. studies & no. of obs.:** 20 studies; 29 indep. obs. for 23 airports. **Meta-regressions estimated:** OLS w/ White se; WLS w/ inverse variance wts.; WLS w/ inverse se wts. **Score:** 23 |
| Rosenberger & Loomis (2000a, b); multi-use recreation | **Effect-size:** CS per activity day; SP & TC studies, so valuation varies. **No. studies & no. of obs.:** 131 studies; 682 obs. **Meta-regressions estimated:** OLS w/ Newey-West standard errors for full model & reduced model; FE & RE panel models. **Score:** 11 |
| Smith & Osborne (1996); stated preference methods | **Effect-size:** WTP visibility improvements at U.S. parks; CV surveys. **No. studies & no. of obs.:** 5 studies; 115 obs. for six parks. **Meta-regressions estimated:** OLS w/ Huber-White standard errors; Box-Cox model; feasible GLS w/ no. of obs. as wts. **Score:** 11 |
| Van Houtven et al. (2007); water quality | **Effect-size:** WTP for water quality; SP studies. **No. studies & no. of obs.:** 18 studies; 131 obs. **Meta-regressions estimated:** WLS w/ Huber-White standard error & wts. based on sample size & no. of obs.; RE panel model. **Score:** 17 |
| van Kooten et al. (2004); global warming | **Effect-size:** Costs of creating carbon offsets using forestry. **No. studies & no. of obs.:** 43 studies; 713 obs. **Meta-regressions estimated:** FE panel model. **Score:** 12 |
| Viscusi & Aldy (2003); value of a statistical life | **Effect-size:** VSL; hedonic wage studies for U.S. & non-U.S. **No. studies & no. of obs.:** 44-46 studies; one obs. per study. **Meta-regressions estimated:** OLS; OLS w/ White standard errors; WLS w/ Huber wts. **Score:** 14 |
| Zamparini & Reggiani (2007a); value travel time | **Effect-size:** Value of travel time as pct of wage; TC surveys. **No. studies & no. of obs.:** 53 studies; 90 obs.; 4 modes of travel. **Meta-regression estimated:** OLS. **Score:** 7 |

Abbreviations: CS=consumer surplus; CV=contingent valuation; FE=fixed effects; OLS=ordinary least squares; RE=random effects; se=standard error; SP=stated preference; TC=travel cost; VSL=value of statistical life; WLS=weighted least-squares; WTP=willingness-to-pay. See the text for discussion of the completeness rating score.