Environmental and Resource Economics **22:** 271–296, 2002.

Is Meta-Analysis a Noah's Ark for Non-Market Valuation?

V. KERRY SMITH[1]* and SUBHRENDU K. PATTANAYAK[2]
[1]*Department of Agricultural and Resource Economics, North Carolina State University Raleigh, NC, USA and Resources for the Future, Washington, D.C. (*author for correspondence, e-mail: kerry_smith@ncsu.edu)* [2]*RII International Research Triangle Park, NC, USA and Department of Forestry North Carolina State University Raleigh, NC, USA*

Abstract. This paper describes meta-analytical methods as they have been applied to non-market valuation research. These studies have been used to review and synthesize literature and, more recently, in benefit transfer. This second use imposes a higher standard on the consistency in economic concepts being summarized and in the resources included in a meta-analysis. To meet this need, the paper proposes and illustrates a structural framework using a generalized method of moments estimator to estimate the parameters of a preference function with the benefits estimates usually encountered in meta-analytic summaries.

Key words: benefit transfer, meta-analysis

1. Introduction

> The advent of micro surveys coupled with the introduction of the computer and the development and dissemination of multiple regression methods by Theil (1961, 1971) and Goldberger (1964) made it possible to produce hundreds, if not thousands, of regressions quickly. The resulting *flood of numbers* was difficult to interpret or to use to test theories or create an informed policy consensus. A demand for low dimensional economically interpretable models to summarize the growing mountains of micro data was created . . . (Heckman 2001: 677, emphasis added)

Over the past twenty-five years micro-econometric analyses of consumer preferences' for environmental amenities have expanded at a dramatic pace. Nonetheless, until recently, the volume of empirical research in environmental economics lagged behind other areas of microeconomics referred to in Heckman's (2001) Nobel lecture.[1] Moreover, in environmental economics the "flood of numbers" has been more diffuse and over a wider range of environmental "commodities." With a clear danger of being overwhelmed, scholars and policy makers in the field have responded by conducting meta-analysis to take stock of what the available, and sometimes competing, empirical estimates imply about the methods being used to estimate the benefits from changes in environmental resources.[2]

The purpose of this paper is to describe meta-analytical methods, especially as applied to non-market valuation, to identify their primary uses, and to interpret the resulting summaries. As the flood of estimates grows, each distinguished by a new methodological twist and each reporting an illustrative benefit computation, we risk losing sight of the goal of the analysis – understanding individual preferences for environmental services.[3] Initially, the primary objective of meta-analyses in environmental economics was to review a literature composed of diverse empirical estimates. Now, the objectives of meta-analyses seem to be more directly linked to policy evaluations, with the meta-regressions being used to generate summaries or benefit transfer functions. While there are clear advantages to imposing the consistent "accounting framework" a meta-regression can provide in summarizing a set of empirical results, there are also important reasons for caution in the uses of these summaries for policy. Based on our review of existing practices and outcomes, we propose a strategy for meta-analysis that focuses on structural models to make meta-analytic summaries consistent with the economic theory underlying environmental valuation.

2. What Is Meta-Analysis?

Meta-analysis refers to the practice of using a collection of formal and informal statistical methods to synthesize the results found in a well-defined class of empirical studies. Glass (1976) first labeled the methods used in these summaries as "meta-analyses" and is usually credited with introducing them into the social sciences.[4] More recently, Cooper and Hedges (1994) describe meta-analysis as a set of methods to synthesize empirical research. Their discussion is within the view of science as a cooperative, cumulative enterprise. For them,

> Theorists provide the blueprints and (empirical) researchers collect the data that are the bricks (p. 4, parenthetical note added).

The result is the "knowledge edifice" comprising a discipline, with meta-analysis improving the building process. While their metaphor helps to illustrate the importance of synthesizing diverse findings, it overlooks the conflict inherent in the development of science. There are certainly complementarities in different aspects of the research contributing to the development of any discipline. However, *there is also competition among ideas*. This competition can take many forms. It promotes efforts to develop new methods to test specific features of existing models and, of course, to propose new models. This process then creates an especially high premium on the ability to develop a consistent, transparent, and reproducible method for drawing conclusions from the diverse array of methods, models and results. Meta-analysis has served this synthesizing role in several disciplines.

Over the past decade, meta-regressions have played a comparable role in summarizing estimates of the economic values for changes in different environmental resources. These meta-analyses have typically served three purposes:

research synthesis, hypothesis testing, and benefit transfer.[5] Before discussing these applications in detail, we outline the stages in a meta-analysis and compare the relative importance of each for summaries of environmental valuation studies versus applications in other disciplines.

Cooper (1982) described research synthesis as itself a research process and identified five stages:

- problem formulation
- data collection
- data evaluation
- analysis and interpretation
- public presentation

The first stage asks: "What evidence should be included in the literature review to learn about a specific question?" To start the process the concept to be measured or summarized must be defined. The analyst must then find variables that correspond to the concept, such as an economic value for a 'water quality improvement,' and distinguish between theoretical and operational definitions. There can be significant selection effects conditioning what is available for a meta-analysis.

Empirical analysis in economics seems to progress through phases – conjecture, confirmation, and closure. After initially postulating a relationship, usually based on theory, early empirical studies are often regarded as speculative. Once the idea being advanced is taken seriously, there is usually scope for a few additional, detailed empirical studies. How few depends on the degree to which each supports theory, the overall confidence in the empirical relationship as dictated by theory, and the policy importance of the empirical issue being studied. Closure arises when a result is "accepted." At this stage, replication rarely finds a home in referred journals.[6]

To be published, non-market valuation research generally must introduce a new method. Field journals in environmental economics are usually not interested in new estimates of the benefits from improving a given environmental resource for their own sake. Updating results for a specific application, such as the demand for sport-fishing recreation or new estimates of the marginal willingness to pay for improvements in air quality may have policy value but usually will not be considered important enough to occupy scarce journal space. A new methodology for dealing with an important issue, such as the opportunity cost of time, in a recreation model is usually required for publication. Similarly, hedonic applications involving effects of a new site-specific environmental attribute (or risk), such as proximity to the route for the transport of nuclear waste (see Kishore and Jenkins-Smith (2001) as an example), are publishable when that study represents the first application, but not necessarily a replication of past work. As a result, new estimators or survey methods applied to "new" environmental resources dominate the published empirical studies. The specific application is simply a vehicle for illustrating the new method. This philosophy stands in sharp contrast to several areas of applied science, where the effects being measured can have inherent

interest and new estimates are published because there is interest in the findings themselves.

The publication philosophy also affects Cooper's stages two and three – data collection and data evaluation. Judgments about what is relevant for a review must consider an important feature of reported results. Often researchers present estimates for economic values as examples, and therefore may report a subset of their findings or emphasize a few of their specific features as they relate to either the new method or the specific application considered. A meta-analysis relying on what is reported may require additional computation. For example, a study reporting on a new econometric approach to estimate a hedonic property value model may include air pollution, but not necessarily report the marginal benefit for an improvement in air quality, i.e., marginal willingness to pay (MWTP) for reductions in air pollution. There are, of course, exceptions but our point is that what is reported is based on the overall objective of each study, which is rarely a simple analysis of the benefits due to changes in the quality or amount of an environmental resource.

Similarly, a conventional travel cost demand model provides a measure of the environmental benefits for changes in price or quality of a recreation site in terms of Marshallian consumer surplus, but not the theoretically appropriate Hicksian measure. Although it is possible to compute the Hicksian measure for a price change from a conventional travel cost model, the computation is not as direct for the case of quality.[7]

Synthesis requires the ability to define a common concept to be measured. As McFadden (1997) noted in a review of two meta-analyses offering evidence on the reliability of contingent valuation studies:

> The core idea of this technique [meta-analysis] is that in an area where there have been multiple statistical studies of the *same phenomenon*, more can be learned by combining information from the separate studies, and the same statistical principles that apply to primary data collection and analysis also apply to this secondary analysis. This last observation is the crux of identifying circumstances under which a meta-analysis can succeed (p. 1, bracketed phrase and emphasis added).

Meta-analyses summarizing non-market valuation studies have often *not* met the goal of measuring "identical" concepts. Consider two ways differences in the phenomenon measured across the studies in a given meta summary might arise. First, the resources being evaluated may be different. For example, recreation demand estimates of consumer surplus per trip may be for different types of recreation sites. In the context of including them in the sample used for a meta-regression, the analyst must consider how these differences are reflected by the independent variables hypothesized to explain the consumer surplus estimates. Often, a set of fixed effects (i.e., one for each type of recreation site) is employed to account for these differences.

Second, the concept being measured, whether Marshallian or Hicksian consumer surplus, may be different. To illustrate the issues raised in assuming that

both consumer surplus measures can be explained with a simple model we can use Willig's (1976) classic paper providing a bound for Marshallian consumer surplus. The income elasticity of demand (η) and the size of the Marshallian surplus (CS) in comparison to income (m_0) were the primary considerations in bounding the difference in the two surplus measures for price changes. One might conjecture that a simple way of reflecting these effects would be to include the mean income reported in each of the studies providing consumer surplus estimates for the meta-regression. Unfortunately, this strategy will not provide the appropriate adjustment. To see why, we can adapt the Willig logic. Suppose we begin our summary of the estimated benefits of price changes by specifying a meta-model for estimates of the compensating variation, as in equation (1).[8] C_j is the estimated compensating variation measure of the surplus for the jth study; x_{ij} (i = 1, 2, ... , k) is a set of variables describing the methodological and resource characteristics for study j; and ε_j an error, reflecting the fact that the function is an approximation.

$$C_j = g(x_{1j}, x_{2j}, \ldots, x_{kj}) + \varepsilon_j \tag{1}$$

To relate Marshallian surplus measures to this model we need to consider the theoretical relationship between the two concepts. Willig (1976) provides a direct expression that can be used to approximate the relationship. It is given in equation (2).

$$C \approx CS + \frac{\eta \cdot (CS)^2}{2m_0} \tag{2}$$

To establish the adjustments to our model in equation (1) when the values used to replace the dependent variable are estimates for Marshallian surplus, we need to solve equation (2) for CS in terms of C and then substitute equation (1) into this expression. Treating equation (2) as a quadratic function and solving for the roots, we have equation (3).

$$CS = \frac{m_0 \pm \sqrt{m_0^2 + 2\eta C m_0}}{\eta} \tag{3}$$

Substituting equation (1) into equation (3) we see that simple adjustments to a meta summary model, such as including income, are unlikely to be effective. At a minimum, if a measure of income from each study in the meta sample is included in the meta-regression, then the parameter reflecting income's influence on the surplus should differ when the dependent variable is a Marshallian or Hicksian surplus. Equation (3) would imply that with large income effects and/or large surplus measures summaries of C and CS should be treated separately.[9]

This example provides one interpretation for McFadden's stated "same phenomenon" requirement. For other benefit measures (such as marginal willingness to pay in Marshallian and Hicksian terms), the requirements will be different. Our point is that when diverse estimates for different resources or from different

methods are included in a single meta-regression an economic framework defining the concepts being measured should be used to evaluate whether the measures can be summarized in a single model.

In some cases, analysts may be willing to admit inclusion of different measures. For example, in the case of hedonic studies most analysts would likely accept summaries where the set of hedonic studies had different measures for the sale price, such as the actual sales price or an average of owners' judgments about what their house would sell for (with appropriate fixed effects to test for differential impacts). They would probably not be willing to assume it is possible to include results based on the construction cost for a new house.

Conventional descriptions of stage 4 of the meta-analytic process identify four types of analysis: vote counting, combining significance levels, combining effects' sizes, and regression summaries. Most meta-analyses associated with environmental values have relied on regression summaries because of their interest in summarizing and explaining the existing results rather than testing a specific hypothesis.

Cooper's (1982) last stage is public presentation of the findings of a research synthesis. Meta-analysis focuses on the conclusions that can be drawn from a review of existing research. There seems to be a presumption in Cooper's description that the analyst developing a meta summary knows in advance which attributes best summarize the set of primary studies. In our experience this has not been the case. A "first pass" summary and resulting multivariate analysis often reveals features of studies we overlooked. For example, coding each individual study as a separate fixed effect may overlook an attribute of a small number of studies that distinguishes how they were done or how the resources studied were described. Study fixed effects would describe such effects as heterogeneity attributed to each author or sample individually, whereas recoding might recognize the common methodological or resource element. Of course, it does not guarantee that the element is the source of the difference, but it does raise this insight as a possibility that can be presented as part of the public presentation of the findings from a meta-analysis.

Many market-based economic applications rely on common databases. When we select a field of inquiry, many studies, especially those involving the macro features of a single economy (e.g., Ricardian equivalence, see Stanley (2001)), use the same data. Thus, a meta-analysis of any differences in findings across studies that use this common data set could focus on the analysts' maintained assumptions.

This is not the case with estimates of environmental values. There are few large-scale public use data sets.[10] Most studies rely on independently collected data. This is especially true for contingent valuation surveys where an important component of the research activity is the design of a set of questions to elicit choice or valuation information that is specific to that particular topic or region. The implication of this feature for environmental applications is that there may well be substantial insights to be gained in these cases by pooling results from different studies.[11]

3. How Has Meta-Analysis Been Used?

Table I provides a selective summary of most of the published and a few of the unpublished meta-analyses involving benefit estimates for changes in environmental resources. It identifies the authors, resource(s) studied, and classifies them based on their primary purpose, including: (a) taking stock of the results based on the relative importance of modeling assumptions and sample characteristics; (b) summarizing the literature to provide a benefits transfer model;[12] and (c) testing a specific hypothesis by exploiting the variation across studies.

3.1. META ANALYSIS TO TAKE STOCK OF PROGRESS

The first applications of meta-analysis to non-market valuation studies, Smith and Kaoru (1990a, b) and Walsh et al. (1990), had different objectives.[13] Smith and Kaoru's work was intended to evaluate the effects of modeling assumptions and data limitations. These authors ask whether there is information ("signal") in the variation ("noise") in consumer surplus estimates across the wide array of travel cost recreation demand studies. They conclude that the character of the data and assumptions influence the results whether measures of consumer surplus per unit of use of the recreation sites (Smith and Kaoru 1990b) or the estimated price elasticity of demand (Smith and Kaoru 1990a). Time constraints, as they affected the opportunity cost of time, the inclusion of cross price or other measures for existence of substitute sites, and the corrections used for the effects of on-site sampling all influenced the consumer surplus estimates.

Walsh et al. (1990) combine travel cost and contingent valuation for resources that support outdoor recreation. Extending the Smith-Kaoru work to include contingent valuation measures per day of use, their study focused on what might be called information transfer rather than a pure assessment of the effects of the research methods used in the empirical studies. In addition to examining the consistency between travel cost and CV estimates, they also sought to investigate whether the cross study comparisons would support "adjustments" to benefit measures as a function of how studies were conducted.[14] Their composite study of travel cost and contingent valuation estimates has been updated to 1998 by Rosenberger and Loomis (2000). Because these authors focus most of their attention on using meta-regressions for benefits transfer, we consider their study in more detail below.

There are several distinctions between Smith and Kaoru's summary in comparison to Walsh, Johnson, and McKean that bear on our evaluation of the other meta-regression summaries. Smith and Kaoru used a single, benefit concept – the Marshallian consumer surplus for access to a recreation site. If this measure was not reported in a study that was part of their sample, they included the study only if it could be computed from the reported information on estimated recreation demand. It was measured relative to the average amount of site use at the average travel cost for each site. By combining travel cost and contingent valuation estimates, Walsh et al. combine Marshallian and Hicksian benefit measures. The authors' objective

Table I. Meta-analyses for non-market valuation studies of environmental resources[a]

Study authors	Year	Environmental resource	Objective considered	Methods	Consistency in concept measured[b]
Smith and Kaoru	1990b	outdoor recreation sites	take stock of literature	travel cost recreation demand	consistent – Marshallian consumer surplus per unit of use
Smith and Kaoru	1990a	outdoor recreation sites	take stock of literature	travel cost recreation demand	consistent - price elasticity of demand
Walsh, Johnson, and McKean	1990	outdoor recreation sites and activities	take stock of literature; evaluate standardization practices for unit values; and benefit transfer	travel cost recreation demand, contingent valuation	inconsistent – without negligible income effects
Smith and Huang	1993	air quality	take stock of literature	hedonic property value	consistent – discrete variable indicating the statistical significance & economic consistency of the effect of particulate matter on housing prices
Boyle, Poe, and Bergstrom	1994	groundwater	take stock of literature, evaluate feasibility for transfer	contingent valuation valuation	consistent – option price for policies intended to avoid groundwater contamination
Smith and Huang	1995	air quality	take stock of literature, evaluate feasibility for transfer	hedonic property value	consistent – marginal willingness to pay for reducing air pollution

Table I. Continued

Study authors	Year	Environmental resource	Objective considered	Methods	Consistency in concept measured[b]
					measured by particulate matter
Smith and Osborne	1996	visibility at Western National Parks	test an economic hypothesis	contingent valuation	consistent – willingness to pay for improvements or avoiding declines in visibility measured by miles of visible range
Loomis and White	1996	threatened and endangered species	take stock of literature, evaluate benefit transfer	contingent valuation	inconsistent
Carson, Flores, Martin, and Wright	1996	environmental resources with revealed preference and stated preference	test hypothesis	travel cost, averting behavior, hedonic property value contingent valuation	inconsistent[c]
Desvousges, Johnson, and Banzhaf	1998	air pollution	benefit transfer	damage function	consistent – measure effect of particulate matter on mortality rate
Brouwer, Langford, Bateman, and Turner	1999	wetland ecosystem	take stock of literature and benefit transfer	contingent valuation	consistent – willingness to pay for wetland services
Rosenberger and Loomis	2000	outdoor recreation sites and activities	take stock of literature and benefit transfer	travel cost recreation demand and contingent valuation travel cost	inconsistent – without negligible income effects

Table I. Continued

Study authors	Year	Environmental resource	Objective considered	Methods	Consistency in concept measured[b]
Markowski, Boyle, Bishop, Larson, and Paterson	2001	recreational fishing sites	take stock of literature, evaluate feasibility for benefit transfer	demand, random utility model, contingent valuation	inconsistent
Shrestha and Loomis	2001	outdoor recreation sites and activities	benefits transfer	travel cost demand, contingent valuation	inconsistent
Woodward and Wui	2001	wetland services	take stock of literature, evaluate for benefit transfer	travel cost demand, contingent valuation hedonic property value, net factor income	inconsistent[d]

[a]In developing the table we attempted to be comprehensive but were limited in our ability to reflect all unpublished meta-analyses. We did omit some meta-analyses, such as Johnson et al. (1997) and Mrozek and Taylor (2001), because they do not deal directly with environmental resource services. The benefit measures summarized are used in conjunction with physical damage functions to measure the benefits from environmental policies. We also omitted meta-analyses that relate to environmental economics but do not deal with benefits such as Baaijens et al. (1998) for regional tourism income multipliers and Cavlovic et al. (2000) for environmental Kuznets curves.

[b]This judgment is made primarily on the basis of the dependent variable used in the meta summary – whether the same concept is being measured across studies. It is a subjective evaluation based on our interpretation, given what is reported in each study. Inconsistent is not intended to mean "incorrect." Rather, it suggests that the summary functions must be evaluated to determine whether they have been or could be adjusted for differences in the concept measured across the studies summarized. These qualifications are important because in some cases authors of the primary studies had as a primary objective comparison of methods that lead to different welfare measures. As a result, use of the results in a meta summary becomes difficult, without attempting to first reconcile these differences. Given the incentives created by criteria for publication, such "methodological" studies are more likely to be the norm rather than the exception.

[c]Carson et al. measure the ratio of a revealed preference to a contingent valuation benefit measure as reported in studies that do comparisons. This will usually involve Marshallian and Hicksian values. In addition, they pool across studies with marginal versus incremental values. These authors' objective was methodological, to provide a "big picture" summary of whether contingent valuation estimates were likely to be substantially different from the revealed preference measures developed in comparison studies. They acknowledge that substantial differences in benefit concept measured, as well as in the goods included in their sample would preclude using their summary for specific conclusions about individual preferences.

[d]Woodward and Wu attempt to reconcile matters with a value per acre; they do face differences in Marshallian and Hicksian values. As a result, this study could be classified either way.

was to adapt what was reported in the primary studies to measure the benefit from a unit of use of the recreation resources involved. Thus, for the cases where travel cost and contingent valuation sought to measure the benefit from current access conditions, these models must maintain the primary studies have negligible income effects.

Other efforts that parallel the early work with recreation include: the Boyle et al. (1994) appraisal of whether eight contingent valuation studies of groundwater protection provide a consistent description of household's valuation of groundwater protection, the Brouwer et al. (1999) and Woodland and Wu (2001) meta-analyses of wetland valuation studies, and the Loomis and White (1996) summary of valuation studies for endangered species.

The Boyle et al. (1994) analysis begins with an explicit formulation of the valuation concept in a groundwater protection CV – an option price. Limitations in both the design of existing studies and what is reported prevents their analysis from accounting for all the core variables describing how groundwater contamination affects people. Instead, they suggest that their summary is the result of a set of compromises. It considers three classes of explanatory variables – the commodity described in each study, respondents' characteristics, and study features. Their conclusions are cautious, noting subjective judgments had to be made about which observations to include, as well as the variables selected to characterize each study's features. In their view, meta-regressions of groundwater values should *not* be used for benefit transfer.

Woodland and Wu provide similar qualifications in their study of wetland values, emphasizing the diversity in the estimated benefit concepts, as well as the overall lack of standardization in estimation methods used and reporting practices across studies. There is an interesting contrast between Woodland and Wu's (2001) meta summary of the willingness to pay for wetland services and the Brouwer et al. (1999) meta summary of studies for the same resources. It highlights the need for adjustments in order to combine benefit estimates from different studies. Woodland and Wu standardize to an annualized value per acre using information on both the size of the area and of the relevant population of users. Woodland and Wu pool benefit estimates across four methods (i.e., net factor input, travel cost, replacement cost, and CV). They reconcile the different valuation concepts to the approximate equivalent of an annualized price for the resource as an asset by converting all values to a per acre format. In contrast, Brouwer et al. (1999) use WTP per household. By confining their summary to CV studies, Brouwer et al. assume that all studies measure a consistent economic concept at the household level, Hicksian willingness to pay for a mix of wetland services. They do note that "... specific WTP questions addressed in each study cover a large continuum of activities, actions or projects related to wetlands, but in some cases (approximately a third of all studies) also to water resources in general" (p. 49). This qualification explains why they include a set of covariates to describe differences in the services being evaluated across studies.[15]

3.2. BENEFITS TRANSFER

An important reason for taking stock of progress in estimating economic values for changes in a particular resource is to use the resulting summary to evaluate policies that modify the services provided by closely related resources. For example, Walsh et al. (1990) developed unit values for recreation activities for use in the benefit analyses conducted as part of the U.S. Forest Service's resource planning process.

Thus, the interest in meta-analytic summaries of existing research relates, in part, to improving the information used in benefit transfers for environmental policy. Research synthesis, when conducted systematically, should help in isolating the "best" of the current information available. This does not mean there is necessarily a defensible measure for the economic value for the change in the quality (or in the amount of a resource) that can be transferred from estimates of other similar resources or that the transfer process is improved using values predicted from meta-regressions. We discuss two studies that are closer to a consistent use of existing evidence and compare them with two that do not.

The two studies offering examples of a consistent transfer synthesize research from benefit measures that are easily defined.[16] The central issue that distinguishes the studies is consistency in both the measure of value summarized across studies and in the environmental commodity or service. The primary research relies on hedonic models and, thus, is a consistent measure of the marginal willingness to pay for a site (or job) specific amenity (or disamenity).[17] Smith and Huang (1995) used the marginal rate of substitution (MRS) between money and air pollution (measured by particulate matter) from hedonic property value models. The reported MRS estimates had to be modified so that they could be used to compute the marginal value of reducing air pollution. Potential economic determinants of the MRS, collected from other sources, were supplemented with variables describing the technical details of how the hedonic model was implemented in each study.[18] This approach accounted for methodological issues while also including the economic determinants of the variations in the MRS across locations.

A second study offering similar prospects for transferable results is the Mrozek-Taylor [forthcoming] summary of the value of statistical life (VSL) results. This meta-analysis combined results from hedonic wage models that included, among other worker and job characteristics, a measure of the risk of a fatal accident on the job. The valuation concept in this application is also a marginal rate of substitution. However, to establish the tradeoff between job risks and wages the model specifies individual well being in terms of expected utility and this is held constant for this definition of the MRS.[19]

The consistency in these two studies is not simply due to their focus on hedonic studies. Rather, it results because these meta-analyses first transform the estimates in the primary studies to a common definition of the benefit to be summarized.[20]

Cases where the meta summaries do not seem to support their use in transfer arise where the valuation concept and/or the commodity valued is not consistent across studies. Table I provides a number of examples. Here, we discuss two in somewhat more detail – Shrestha and Loomis (2001) and Loomis and White (1996). Shrestha and Loomis (2001) use the updated Rosenberger-Loomis (2000) recreation database for an international benefit transfer.[21] The dependent variable is described as consumer surplus per day. The methods in their analysis include: conventional travel cost demand, hedonic travel cost, random utility models (RUM), and contingent valuation surveys. RUM and CV provide Hicksian, while conventional travel cost and hedonic travel cost yield Marshallian measures.

Earlier we presented an example describing how Hicksian and Marshallian measures for a price change could be linked to each other if they are to be included in a single model. When the benefit methods are expanded to include hedonic travel cost (HTC) and RUM, the differences in the economic decision frameworks are more pronounced. Often HTC and RUM focus on describing choices as if they involve single choice occasion.[22] Benefit estimates derived using the travel cost demand and contingent valuation models are usually designed to measure seasonal (or annual) benefits. Simply scaling the first two approaches (i.e., HTC and RUM) by the number of trips for the season would not resolve the differences in the income measure and its potential influence on individual choices. This would suggest that small income effects may be required to use all findings in a consistent meta summary.

Loomis and White (1996) pool results across different endangered species and identify the species class (e.g., bird, fish, marine mammal, with a composite including wolves, sheep, bears, and turtles as the omitted group). In this case, the valuation concept is consistent across studies – a Hicksian surplus measure – but the "commodities" are not. Our discussion of the conceptual issues in pooling results across studies with different commodities identified the question of how different the commodities can be before a joint summary becomes implausible. One way to address this question is to formulate a model of preferences in which each resource contributes to an aggregate index of resource services. If the individual resources contributing to the aggregate are close substitutes for each other, then small changes in any one resource individually will lead to approximately the same willingness to pay (WTP) measures. As the differences in substitution elasticities become more pronounced, so also will the differences in the WTP measures for changes in the amounts of individual resources. Moreover, these WTP measures will depend on the base level of the individual resource from which the change takes place.

When we evaluate the set included in Loomis and White's summary, it seems reasonable to speculate that these resources are quite diverse. Some would be associated with consumptive forms of recreation (i.e., hunting and fishing) and others would not. Some are often sought after in trips to view wildlife while others are

not. Thus, the heterogeneity seems to us quite pronounced. This does not seem to be Loomis and White's prognosis. They are more optimistic that a meta-regression based on these types of composite samples would be ready for benefit transfers in the near future.[23]

One approach for dealing with the effects of heterogeneity in the resources combined in a meta summary is to distinguish between the uses of meta-regression summaries. We would argue that the consistency requirements are higher for benefit transfers, as opposed to meta-regressions for reviewing the literature. These benefit transfers are likely to require greater consistency in the valuation concept (rather than just resource consistency) than one might require in using meta-analytic techniques to synthesize research. Consistency can be achieved with diverse valuation concepts across different methods, provided each can be reconciled with an economic model of the choice process.[24]

3.3. TESTING HYPOTHESES

When a consistent economic concept of value is measured across studies used in a meta-analysis, then in principle, it should be possible to test hypotheses using the cross study variation in conditions. This was McFadden's point and has been one of the primary uses for applications of meta-analysis to medical and educational research. Testing features of individual preferences, based on cross-study differences, requires replication of valuation studies with comparable methods and a focus on modifications in the changes in the resource to be valued. There are few areas in non-market valuation where this has been possible. One of them involves contingent valuation of visibility changes at national parks. Because of concerns to protect air quality in pristine ClassI areas (due to provisions of the Clean Air Act and its amendments) and the rapid growth in areas nearby, there has been interest in measuring the economic value of avoiding visibility changes.

Smith and Osborne (1996) took advantage of this interest and used a set of different studies to test whether CV methods would be judged sensitive to the scope (or size) of the change in an environmental resource. Because the valuation concepts were comparable and the commodity was presented using the same basic format – photos of scenes in Western national parks with differences in the visible range of sight – these studies met both of the consistency standards we identified and a meta-analysis offered an opportunity to test the scope hypothesis.

Meta-analysis of the findings has a further advantage – it allows the results to be used to consider whether there might, as McFadden conjectures, be other attributes that would also contribute to visibility (i.e., cloud cover, time of the day, etc.). By establishing a role for miles of visible range through a meta-analysis, it is possible to consider the required design points, with variations in these other attributes contributing to visibility, that would be necessary to establish their relationship to changes in the visible range.

4. Structural Meta-Analysis

Our concerns with using meta-regression summaries for benefits transfer and in testing economic hypotheses stem from inconsistencies in the commodities being evaluated and in the economic concepts being measured across the studies in each type of meta summary. To date, analysts have selected one of three responses to these inconsistencies:

(a) include all measures without adjustment to a common economic concept and attempt to control for differences in commodities with covariates;

(b) adjust the measures to a common economic concept using information from the primary study or from other sources that attempt to match the original study, and only use those studies where common measures have been developed;

(c) drop studies that don't fit some standard for comparable commodities and consistent measures.

None of these approaches is ideal. However, as our earlier comments imply, we recommend the second and third approaches over the first. In what follows we focus our attention on a proposal to improve the second strategy.

An important limitation to this strategy is that the adjustments are "behind-the-scenes" and users of the resulting summaries must accept maintained hypotheses linking the economic concept in each primary study to actual measure reported, and then use this reported measure in their meta-analytic summary. For example, in the case of a hedonic property value study, the primary hedonic price equation might be a log linear specification with estimated parameters interpreted as elasticities. The marginal rate of substitution (or marginal willingness to pay) equals the marginal price (i.e., the change in the price with respect to the attribute of interest). In the case of a log-linear model, additional data on the levels of the price and the attribute must be used to compute the marginal price from the estimated parameters for the model. For travel cost models reporting demand functions with Marshallian consumer surplus per trip as the intended summary concept, the transformations can be more extensive.

One way to address concerns about the judgments required to adjust benefit measures across studies is to place at the center of the analysis the structural model that underlies all the benefit measures and use it to inform the adjustment process. Under what we are calling the "structural meta-analysis" the analyst specifies a preference function as the structural model and uses it to interpret estimates across different valuation studies. If one study measures a marginal willingness to pay at some level of environmental resource and another study measures the willingness to pay for a specific change in the same environmental resource, one can define each benefit measure using a common preference function. Each benefit measure can then be expressed as a specific function of the underlying preference parameters. The task of a structural meta-analysis is to identify and estimate the parameters of the preference function.

In this setting the issues of benefit transfer or hypothesis testing can be addressed based on the ability to estimate of the preference parameters. One early stumbling block for this strategy was the desire to treat the stochastic errors associated with different behavioral models consistently. It was envisioned that this process would entail beginning with a preference specification and incorporating errors (as unobserved heterogeneity) that entered in the resulting estimating models in ways that permitted maximum likelihood estimation. Cameron et al.'s (2002) comparison of different non-market value elicitation methods offers an example of the implicit logic. Because each approach can be consistently represented within a given preference and error specification, it is possible to define a composite likelihood function. Unfortunately, when analysis must consider different methods, margins for observing choices, and a wide array of investigators, this goal is not as easily attained.

To avoid these issues we used a generalized method of moments approach, by recognizing that each benefit measure can be expressed as a moment equation that links it to the preference function. To illustrate the logic we selected CV estimates of the willingness to pay for air pollution reductions in Los Angeles (reported in Brookshire et al. (1981, 1982)) and hedonic property value estimates of the incremental value of visibility improvements in the same area (reported in Beron et al. (2001)).[25] Both studies convert air pollution changes to the same concept – changes in visible range – *and* both studies consider the same area.

The structural approach for summarizing benefit estimates requires the specification of a preference function and the definition of each concept measured in each of the studies in terms of that function to be included in a meta-summary. Equation (4) is the specification we selected for the indirect utility function.[26]

$$V = (\theta_1 v)^b + \sum_{j=2}^{K} (\theta_j A_j)^b + \left[p^{-\alpha}(m - R(v, A_2, \ldots, A_K))\right]^b \tag{4}$$

v is the measure of air pollution, expressed as visibility in miles of visual range; $A_2, A_3, \ldots, A_K$ are the housing and location attributes; m corresponds to household income; $R(.)$ is the hedonic price function (expressed as an annual rent); and P is a composite price index for the other goods consumed by the household.[27] α, b, and $\theta_1, \ldots, \theta_k$ are the unknown structural parameters. The Brookshire et al. (1981, 1982)] CV studies provide a willingness to pay for a specific change in visibility (described using photos of visibility change) and the Beron et al. (2001) valuation study provides estimates of marginal willingness to pay for an incremental change in visibility evaluated at a specified level.

To derive an estimating equation for the CV estimates, consider the definition of the WTP of a visibility change from v^0 to v^1 (with $v^1 > v^0$) as in equation (5).

$$(\theta_1, v^1)^b + \sum_{j=2}^{K} (\theta_j A_j)^b + [P^{-\alpha}(m - R(\cdot) - WTP)]^b$$
$$= (\theta_1, v^0)^b + \sum_{j=2}^{K} (\theta_j A_j)^b + [P^{-\alpha}(m - R(\cdot))]^b \quad (5)$$

In this expression we assumed that the visibility change was not capitalized in the rent; that is, the respondents to the Brookshire et al. CV survey did not perceive this connection. We could alter this assumption. It is not a requirement of the method. Simplifying equation (5) and solving for WTP yields the first of our estimating equations, given as (6a) below. Notice that all the other housing attributes that are assumed to remain unaffected by the visibility change cancel in the definition of the estimating equation. This simplification is deliberate and results from the preference specification we adopted.

The expression for the marginal willingness to pay in terms of the marginal price can also be derived from equation (4) by taking the total differential, assuming utility is maximized at each individual's current location (i.e., $dV = 0$). In this case, we consider variations in the rent with v because this feature is what will be observed through the hedonic models. Simplifying the total differential of equation (4) yields the estimating equation (6b).

$$WTP = (m - R(\cdot)) - \left[(\theta_1 v^0)^b - (\theta_1 v^1)^b + [(m - R(\cdot))]^b\right]^{1/b} \quad (6a)$$

$$\frac{\partial R}{\partial v} = \frac{\theta_1^b (v^3)^{b-1}}{(m - R(\cdot))^{b-1}} \quad (6b)$$

With estimates of the incremental WTP and the marginal price, then equations (6a) and (6b) provide the moment conditions for our estimates of the structural parameters θ_1 and b. By selecting the same geographic area we assume there is no variation in P across the households in the two samples. As a result, we cannot identify a basis for estimating α and the terms involving P have been deleted from (6a) and (6b) by normalizing P to unity.

Three aspects of these estimating equations are important to the implementation of meta-analytic summary. First, this framework assigns specific roles for the attributes of each type of study. For example, with (6a) both the baseline and new visibility levels (and not just the change in visibility) should enter the estimating equation. Moreover, if these WTP estimates are to be pooled with hedonic measures, income should be measured as net income (i.e., income less annual rental for housing). Second, the level of visibility in the two equations need not overlap. v^3 represents the level used to compute the marginal prices, while v^0 and v^1 are the levels in the CV questions. Finally, to use these equations as moments we add errors (i.e., U_1 and U_2 for equations (6a) and (6b), respectively) and consider the

moments of each with a set of instrument variables (Z). Equations (7a) and (7b) illustrate the logic.

$$E(Z^T U_1) = \\ E(Z^T(WTP - (m - R(\cdot)) + [(\theta_1 r^0)^b - (\theta_1 r^1)^b + (m - R(\cdot))^b]^{1/b})) = 0 \quad (7a)$$

$$E(Z^T U_2) = E\left(Z^T\left(\frac{\partial R}{\partial v} - \frac{\theta_1^b (v^3)^{b-1}}{(m - R(\cdot))^{b-1}}\right)\right) = 0 \quad (7b)$$

where Z is a vector of instruments

To estimate the parameters the two equations are stacked and a restricted distance estimator is used. That is, θ_1 and b are restricted to be the same in the two equations. This constraint recognizes the fact that these parameters appear in both equations and link back to the benefit function for transfer in (6a) and to the preference function in (4).

Our application has 72 estimates of WTP for different visibility changes from the Brookshire et al. study. Beron et al. (2001) provided 36,160 estimates of the marginal willingness to pay (MWTP), qualitative variables associated with the county identification (Los Angeles, Orange, Riverside, or San Bernadino), average income for the census tract (net of annualized housing price), and percent with a college degree in the census tract relevant to each housing sale.

To illustrate our approach we randomly sampled 72 observations from the Beron et al. hedonic estimates of marginal willingness to pay ten times and used each sample from this process together with the 72 CV estimates from the CV study.[28] All estimates were converted to 1995 dollars. The 1994–95 Consumer Expenditure Survey for western metropolitan statistical areas was used to estimate the fraction of income devoted to housing expenditures, the concept used in equation (6a).

Table II reports the joint estimates for the ten randomly drawn samples from the Beron et al. database. Each was used with the same sample of CV results to compute the method of moments estimates of the structural parameters that can be identified with these measures.[30] The numbers in parentheses report the estimated asymptotic standard errors for each parameter. The last three columns provide the mean values for the key variables used from the Beron et al. study to estimate the marginal willingness to pay moment condition.

It seems clear that b is more precisely estimated than θ_1 and is also more stable across different samples. Both contribute to the willingness to pay function in equation (6a), which would provide a basis for structural transfers. This strategy estimates the parameters needed to evaluate the benefits of visibility change via benefits transfer, provided other housing attributes do not change and the set of prices is comparable between the areas used for the estimation and transfer sites.

However, these estimates are not ready for use in a benefits transfer. What we summarized is an early report on a more extensive effort to refine the method.[30] There are a number of problems to overcome. With the limited instruments available, convergence was often difficult to realize. We expect that with refinement in

Table II. A comparison of effects of sampling on the structural meta-analytic estimates[a]

	θ_1	b	v^3	m-R(·)	MWTP
Sample 1	1.982	0.836	15.33	43174	597.06
	(2.849)	(0.172)			
Sample 2	1.221	0.792	15.29	46897	623.64
	(2.168)	(0.188)			
Sample 3	2.790	0.907	15.56	41038	502.77
	(4.075)	(0.201)			
Sample 4	1.885	0.821	15.05	48801	617.61
	(2.611)	(0.161)			
Sample 5	2.010	0.843	15.89	43114	550.69
	(3.001)	(0.181)			
Sample 6	1.827	0.817	15.39	45782	563.42
	(2.558)	(0.162)			
Sample 7	1.781	0.826	15.61	42562	559.98
	(2.666)	(0.173)			
Sample 8	1.745	0.835	15.24	44210	545.84
	(2.775)	(0.186)			
Sample 9	1.216	0.796	15.34	46554	529.76
	(2.180)	(0.190)			
Sample 10	2.558	0.842	15.40	42162	574.06
	(2.945)	(0.144)			

[a]The numbers in parentheses below the estimated parameters are the estimated asymptotic standard errors. The last three columns report the mean values from the random samples of 72 observations taken from the Beron et al. (2001) computations. They report the average of the visibility values in miles used to compute the marginal willingness to pay from the hedonic price functions (MWTP) for visibility and the average household income net of housing costs (*m-R(.)*) in 1995 dollars.

the set of instruments and expansion in the sample, the precision of our estimates of θ_1 can be enhanced and convergence properties are also likely to be improved.

We are in the process of considering several extensions to this framework to attempt to address these limitations. First, we could use different sized samples for the two conditions by appealing to the logic used in combining multiple micro-sample or micro and macro-samples (see Angrist and Krueger (1992) and Imbens and Lancaster (1994) as examples). Second, we could choose studies from different areas. Extending the CV or hedonic studies to other areas would offer a basis for observing variation in P and, thus, a basis for identifying estimates of α. Third, if we are willing to specify how other aspects of air pollution affect preferences, we would not be limited to studies associated with measuring the economic benefits

from visibility improvements. Finally, we need not overlook adjustments for the methods used, as determinants of the findings (e.g., valuation method type), in implementing this moments method. This structural approach allows these methodological factors to be considered as fixed effects, shifting the moment conditions associated with the benefit measurement method where they arise.

More generally, the approach could be applied to combining travel cost and CV as well as other methods. In many respects, it parallels the logic proposed by Cameron (1992) for joint estimation of stated and revealed preference responses from a single sample. The difference is that in our case we are using a structural model to provide a basis for a meta-analytic summary of diverse types of nonmarket valuation estimates.

5. Launching the Ark

Most meta-analyses seek to review and synthesize an extensive literature, often with diverse findings. These statistical summaries can help to understand the reasons for that diversity. In some cases there are sufficient replications with the same concept measured and environmental resource studied that it has been possible to test economic hypotheses. These meta-analytic studies from nonmarket valuation research parallel the application of meta-analysis in other fields. The use of meta-analytic summaries in benefit transfer for policy seems to be new to environmental applications. We have argued it imposes a higher standard on assuring that both the environmental resources considered and the economic concepts being measured are consistent across the studies combined in a meta summary intended for benefit transfer.

Heckman (2001: 678) notes that with structural models "empirical knowledge can be cumulated consistently." His call for low dimensional (in terms of parameters) models focuses our attention on what is learned about human behavior in the context of environmental resources and the clues they offer about environmental values. In this paper, we have proposed that meta-analysis can be a form of a low dimensional model if we use economic theory in the form of a preference structure to guide the meta-analysis and the cumulative learning process. Although the learning would occur in terms of better estimates of preference parameters, we do not argue that all preference parameters must be identified. Instead, by consistently using the available information our proposed approach ensures that we can identify the subset of the structural parameters required for a benefit transfer associated with the environmental resource involved. This process also recasts the problem of benefit transfer and provides a basis for judging whether there is sufficient research to inform specific types of policy questions.

Our Noah's Ark metaphor was intended to portray a situation where not only are we saved from a disastrous flood, but it is possible to maintain the integrity of a unifying framework for understanding the "flood of numbers."[31] In our view, meta-analysis, as it has been applied to date, has not yet fully realized these dual

objectives. Low dimensional structural summaries are essential to meeting those goals. However, we need further research evaluating the feasibility and importance of using a single, structural framework to summarize what we know about people's preferences for environmental goods. The floodwaters are rising, but the meta-analyses of non-market valuation studies for environmental resources are growing in number and influence.

Acknowledgements

Partial support for Smith's research was provided by the U.S. Environmental Protection Agency under a grant (R-826609-01) and a cooperative agreement with the Office of Water. Partial support for Pattanayak's research was provided by USDA Forest Service (Economics Research Unit SRS-01-CA-11330143-441) and U.S. Environmental Protection Agency (Office of Water 68-C-01-142). However, the views reported here do not reflect the opinions of either agency. Thanks are due to Richard Carson and two anonymous reviewers for very helpful comments on an earlier draft and to Hyun Kim for research assistance with the computations associated with the method of moments estimator, Jim Murdoch for assistance in obtaining some of the data used in that application, and Susan Hinton for producing coherent versions of this and several earlier drafts at a record pace.

Notes

1. See Carson (forthcoming). This summary indicates that the contingent valuation literature has expanded from slightly over 2,000 studies at the beginning of 1995 to substantially over 5,000 studies as of the end of 2001.
2. Stanley (2001) describes meta-analytic methods as offering a means to increase the effectiveness of literature reviews. In a description quite similar to Heckman's commentary, he notes that:

   > In an era characterized by the expansion of research publications and an avalanche of information, balanced and critical literature reviews serve a crucial function. They act as intelligent agents searching through mountains of potentially contradictory research to uncover the nuggets of knowledge that lie buried underneath (p. 131).

   For him, meta-analysis helps to make the process more systematic and to avoid bias in these reviews. We argue in what follows that it can do more, especially in non-market valuation where independent data sets are available to evaluate consumer preferences for environmental resources.
3. In practice, the demand function we seek to estimate (and the implied values for environmental goods) is for changes probably best described with a Hicksian virtual price function (an income composition function.
4. Cooper and Hedges (1994) cite several early examples of efforts to estimate effect sizes by combining results from separate samples. Their examples date to 1904 (see pp. 5–6). Moreover, they also highlight the parallels in statistics to combining probabilities across studies (or independent tests in the same study).
5. See Cooper and Hedges (1994) for a summary and elaboration of Cooper's description of the characteristics of each stage in their Table 1.2. Most of this discussion relates to situations where there are less formal models than we use to describe behavior in micro-economic applications.

6. This pattern leading to an "end" of professional interest in a topic for empirical research may not be a problem for some types of economic applications. However, it is a serious problem for the estimates of people's economic values for changes in environmental resources. These estimates are only partial descriptions of people's preferences for each resource. As a result, these values are not unknown constants. They depend on the constraints, available substitutes, and other factors influencing people's choices. If these features change, then the measured economic values for the same change in an environmental resource would change. The important policy role for measures of economic value implies they need to be periodically updated to reflect current conditions.
7. See Bockstael and McConnell (1993), Larson (1991), and Smith and Banzhaf (2001) for further discussion.
8. A price change can represent the loss of a resource providing use-related recreation by assuming the price changes from the current conditions which permit use to the choke price which is defined by the fact that it is high enough that an individual would choose no consumption of the good involved.
9. A similar argument to this one could be developed for the three consumer surplus concepts associated with quality or quantity changes by using Hanemann (1991) and adapting our earlier logic.
10. Notable recent exceptions in the U.S. include the 1994 National Survey of Recreation and the Environment which was conducted in two separate formats – one sponsored by the Economic Research Service of USDA and a second by EPA. The ERS survey was organized around four sub-state areas with intensive agriculture and the EPA was intended to be a representative sample of the U.S. population with sufficient information to measure water based recreation. A new outdoor recreation survey for 2000 was recently completed. Other efforts in the U.S. through the U.S. Forest Service and the National Oceanic and Atmospheric Administration (NOAA) have involved recreation sites as well as more targeted environmental resources, such as the Florida Keys.
11. Of course, the pooling across studies to compose the sample must also recognize the need for consistency in the economic concept of benefits used in each of the primary studies as our earlier discussion suggested.
12. Benefits transfer refer to the practice of adapting estimates of the economic value of changes in the access conditions or quality of resources from existing research and using them, usually in some modified form, to assess the economic value of a similar but separate change in a different resource.
13. Nelson (1980) used a basic form of meta-analysis when he summarized evidence on the relationship between airport noise and property values. However, his analysis focused on putting measures of the effects of noise on property values in comparable terms and constructing a cross study mean as an estimate of the overall effect. This research did not attempt to use statistical methods to explain differences in the effects attributed to noise across each primary study. Thanks are due to Raymond J.G.M. Florax for bringing this study to our attention.
14. Most of these adjustments focused on travel cost demand studies when used to measure consumer surplus per day. The context for this aspect of the work stems for the U.S. Forest Service's need for "unit day values" to measure the benefits from providing resources that support outdoor recreation. Prior to the Walsh, Johnson, and McKean work, conventional practice has called for adjusting existing benefit studies so they were in constant dollars and measured per standardized activity day. In addition, Loomis and Sorg (1984) had proposed a series of adjustments to these unit benefits based on modeling judgments. These included: (a) increasing reported travel cost consumer surplus estimates by 30% for the omission of travel time, (b) increasing both travel cost and contingent valuation by 15% for the omission of out-of-state users, and (c) decreasing travel cost models' estimates by 15% when individual observations were used. An important aspect of the Walsh, Johnson, and McKean work involved

comparing the differences across studies with differences in attributes with these proposed adjustments.

15. Smith (1997) argues that it might be useful to distinguish the plan, project, or policy presented in CV studies as the method to be used to change the environmental resource as the commodity. This is what people decide about. To use the results from their decisions, analysts are implicitly making assumptions about the extent of substitution between the plan and the change in the environmental resource of interest.
16. Since the first author was involved in one of these studies, this judgment can hardly be considered a detached assessment.
17. Hedonic models rely on a market equilibrium to reveal the marginal tradeoffs for attributes available exclusively through a location (in the case of property value models) or a job (in the case of wage models). The primary assumptions of the model are: the existence of the equilibrium (described by the price or wage function); the presence of a sufficient diversity of choices to allow an approximately continuous price locus; full knowledge of the market tradeoffs among all participants; and inability of any one agent to influence what is offered or paid for through increments in the attributes.
18. Each study was augmented with information about the overall air quality and housing market conditions, as well as the economic features of the households (e.g., income) in the year and city of each hedonic analysis.
19. Their analysis does not attempt to characterize fully the labor market conditions or the individuals involved in the choices being described in their models. The authors do, however, include measures of the national unemployment rate in the year the study was conducted and qualitative descriptions identifying whether it was a national or local labor market.
20. Other examples aside from hedonic studies where the meta summaries defined a single consistent basis for their analysis would include the Desvousges et al. (1998) summary of concentration/response functions as well as their analysis of willingness to pay to avoid morbidity effects.
21. The authors offer a positive (with acknowledged wide error bands) view of their meta-regression based benefits transfer, noting that:

    > From paired t-tests of means of meta-predicted values and means of original CS (consumer surplus per day) values obtained from the ROW (rest of world) studies, it is identified that predicted values using a meta model estimated for the United States were significantly different from the original study value for up to half of the cases ... However, the absolute average error was only 28% across all computations (30% using means of the method variables), an often acceptable level in most benefit transfer settings (pp. 80–81).

22. The HTC can be a seasonal model. However, often to avoid addressing the tasks posed in adding choice-occasion levels of attributes across trips, analysis has been confined to the decision on one trip.
23. The authors do suggest that at best such estimates would "... provide a rough first estimate to determine whether the benefits are likely to be much larger or much smaller than the costs" (p. 204).
24. A separate issue that arises with using meta summaries for benefits transfer concerns the appropriate treatment of methodological variables (e.g., functional form used, estimator, etc.). When the valuation concept is consistent, the recommended values for these method-based differences are not clear. One interpretation is to use them to adjust results for the set of assumptions regarded as most plausible. This strategy is the logic adopted in Walsh et al., Loomis and White, and Mrozek and Taylor. Another interpretation is that by controlling for method based differences in study results we can attribute differences in results to the various sources estimation uncertainty (i.e., the maintained assumptions with each method). In this case we would want to use the meta summary to minimize the effect of each method on the results used in a transfer.

25. We would not have been able to develop these results without James Murdoch's willingness to share detailed estimates of the marginal willingness to pay results from Beron, Murdoch and Thayer (2001). We are very grateful to Jim and his co-authors for their help.
26. See Smith et al. (2002) for further discussion of its properties.
27. Income, annualized housing price, and the composite price index are assumed to be normalized by the price of a numeraire which is invariant across households to assure homogeneity of degree zero in income and prices.
28. This strategy avoids the weighting issue that necessarily arises with such large differences in the two study samples. See Angrist and Krueger (1992) and Imbens and Lancaster (1994) for discussion of related issues.
29. Our instruments (i.e., the Z's) for the marginal willingness to pay included the percent college in the tract, the county dummy variables, level of visibility, income, and rent and for the incremental willingness to pay the characteristics of the CV questions including: whether the change was described as a decline or an increase, whether it was described as certain or not, and the level of initial and final visibility. The generalized method of moments estimator was developed using LIMDEP, Version 7; see pp. 401–402 of the manual for a description of the weighting function used in this program's GMM estimator.
30. The more extensive research evaluates the treatment of sample size, the number of moment conditions used and expands the set of instruments included in the estimation. It is being undertaken with George van Houtven.
31. Weitzman (1998) has also used the Noah's Ark metaphor in his paper considering how to preserve diversity under a limited budget constraint. Thus, parsimony as reflected in his fundamental ingredients for choosing among alternatives to preserve has a value in this case as well.

References

Angrist, J.D. and A.B. Krueger (1992), 'The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments rom Two Samples', *Journal of the American Statistical Association* **87** (June), 328–336.

Baaijens, S. R., P. Nijkamp and K. V. Montfort (1998), 'Explanatory Meta-Analysis for the Comparison and Transfer of Regional Tourism Income Multipliers', *Regional Studies* **32**(9), 839–849.

Beron, K., J. Murdoch and M. Thayer (2001), 'The Benefits of Visibility Improvement: New Evidence from the Los Angeles Metropolitan Area', *Journal of Real Estate Finance and Economics* **22** (March/May), 319–338.

Bockstael, N. E. and K. E. McConnell (1993), 'Pubic Goods and Characteristics of Non-Market Commodities', *Economic Journal* **103** (September), 1244–1257.

Boyle, K. J., G. L. Poe, and J. C. Bergstrom (1994), 'What Do We Know About Groundwater Values? Preliminary Implications from a Meta Analysis of Contingent-Valuation Surveys', *American Journal of Agricultural Economics* **76** (December), 1055–1061.

Brookshire, D. S., R. C. d'Arge, W. D. Schulze and M. A. Thayer (1981), 'Experiments in Valuing Public Goods', in V. Kerry Smith, ed., *Advances in Applied Microeconomics*, Vol. I. Greenwich, CT: JAI Press, pp. 123–172.

Brookshire, D. S., M. A. Thayer, W. D. Schulze, and R. C. d'Arge (1982), 'Valuing Public Goods: A Comparison of Survey and Hedonic Approaches', *American Economic Review* **72**(1), 165–177.

Brouwer, R., I. Langford, I. Bateman and R. K. Turner (1999), 'A Meta-Analysis of Wetland Contingent Valuation Studies', *Regional Environmental Change* **1** (November), 47–57.

Cameron, T. A. (1992), 'Combining Contingent Valuation and travel Cost Data for the Valuation of Nonmarket Goods', *Land Economics* **68**(3), 302–317.

Cameron, T. A., G. L. Poe, R. G. Ethier and W. D. Schulze (2002), 'Alternative Non-market Value-Elicitation Methods: Are the Underlying Preferences the Same?' *Journal of Environmental Economics and Management* (in press).

Carson, R. T., N. E. Flores, K. M. Martin and J. L. Wright (1996), 'Contingent Valuation and Revealed Preference Methodologies: Comparing the Estimates for Quasi-Public Goods', *Land Economics* **72**(1), 80–99.

Carson, R. T. (forthcoming), *Contingent Valuation: A Comprehensive Bibliography and History*. Northampton, MA: Edward Elgar.

Cavlovic, T. A., K. H. Baker, R. P. Berrens and K. Gawande (2000), 'A Meta-Analysis of Environmental Kuznets Curve Studies', *Agriculture and Resource Economics* **29**, 32–42.

Cooper, H. M. (1982), 'Scientific Guidelines for Conducting Integrative Research Reviews', *Review of Educational Research* **52**, 291–302.

Cooper, H. M. (1988), 'Organizing Knowledge Synthesis: A Taxonomy of Literature Reviews', *Knowledge in Society* **1**, 104–126.

Cooper, H. and L. V. Hedges (1994), 'Research Synthesis as a Scientific Enterprise', in H. Cooper and L. V. Hedges, eds., *The Handbook of Research Synthesis*. New York: Russell Sage Foundation, pp. 3–14.

Desvousges, W. H., F. R. Johnson and H. S. Banzhaf (1998), *Environmental Analysis with Limited Information*. Cheltenham, UK: Edward Elgar.

Glass, G. V. (1976), 'Primary, Secondary, and Meta-Analysis', *Educational Researcher* **5**, 3–8.

Hanemann, W. M. (1991), 'Willingness to Pay and Willingness to Accept: How Much Can They Differ?' *American Economic Review* **81** (September), 635–647.

Heckman, J. J. (2001), 'Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture', *Journal of Political Economy* **109**(4), 673–748.

Imbens, G. W. and T. Lancaster (1994), 'Combining Micro and Macro Data in Microeconometric Models', *Review of Economic Statistics* **61** (October), 655–680.

Johnson, R., E. Fries and S. Banzhaf (1997), 'Valuing Morbidity: An Integration of the Willingness-to-Pay and Health Status Index Literature', *Journal of Health Economics* **16**, 641–645.

Kishore, G. and H. Jenkins-Smith (2001), 'Nuclear Waste Transport and Residential Property Values: Estimating the Effects of Perceived Risks', *Journal of Environmental Economics and Management* **42** (September), 207–233.

Larson, D. M. (1991), 'Recovering Weakly Complementary Preferences', *Journal of Environmental Economics and Management* **21** (September), 97–108.

Loomis, J. B. and D. S. White (1996), 'Economic Benefits of Rare and Endangered Species: Summary and Meta-Analysis', *Ecological Economics* **18**, 197–206.

Markowski, M. A., K. J. Boyle, R. C. Bishop, D. M. Larson and R. W. Paterson (2001), 'A Cautionary Note on Interpreting Meta Analyses', unpublished paper, Industrial Economics Inc.

McFadden, D. (1997), 'Can Meta-Analyses of CV Studies Determine Their Reliability?', unpublished paper, Department of Economics, University of California, Berkeley.

Nelson, J. P. (1980), 'Airports and Property Values: A Survey of Recent Evidence', *Journal of Transport Economics and Policy* **19**, 37–52.

Rosenberger, R. S. and J. B. Loomis (2000a), 'Using Meta-Analysis for Benefit Transfer: In-sample Convergent Validity Tests of an Outdoor Recreation Database', *Water Resources Research* **36** (April), 1097–1107.

Rosenberger, R. S. and J. B. Loomis (2000b), 'Panel Stratification in Meta-Analysis of Economic Studies: An Investigation of Its Effects in the Recreation Valuation Literature', *Journal of Agricultural and Applied Economics* **32**(3), 459–470.

Shrestha, R. K. and J. B. Loomis (2001), 'Testing a Meta-Analysis Model for Benefit Transfer in International Outdoor Recreation', *Ecological Economics* **39**, 67–83.

Smith, V. K. (1997), 'Pricing What is Priceless: A Status Report on Non-market Valuation of Environmental Resources', in H. Folmer and T. Tietenberg, eds., *The International Yearbook*

of Environmental and Resource Economics 1997/1998: A Survey of Current Issues. Cheltenham, UK: Edward Elgar, pp. 156–204.

Smith, V. K. and S. Banzhaf (2001), 'A Diagrammatic Exposition of Weak Complementarity and the Willig Condition', working paper, Center for Environmental and Resource Economic Policy, North Carolina State University, October.

Smith, V. K. and J-C. Huang (1993), 'Hedonic Models and Air Pollution: Twenty-Five Years and Counting', *Environmental and Resource Economics* **3** (August), 381–394.

Smith, V. K. and J-C. Huang (1995), 'Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models', *Journal of Political Economy* **103** (February), 209–227.

Smith, V. K. and Y. Kaoru (1990a), 'What Have We Learned since Hotelling's Letter? A Meta-Analysis', *Economics Letters* **32** (March), 267–272.

Smith, V. K. and Y. Kaoru (1990b), 'Signals or Noise? Explaining the Variation in Recreation Benefit Estimates', *American Journal of Agricultural Economics* (May), 419–433.

Smith, V. K. and L. L. Osborne (1996), 'Do Contingent Valuation Estimates Pass a 'Scope' Test? A Meta-analysis', *Journal of Environmental Economics and Management* **31** (November), 287–301.

Smith, V. K., G. van Houtven and S. K. Pattanayak (2002), 'Benefit Transfer via Preference Calibration: Prudential Algebra for Policy', *Land Economics* **78** (February), 132–152.

Sorg, Cindy F. and John B. Loomis (1984), 'Empirical Estimates of Amenity Forest Values: A Comparative Review', General Technical Report, RM-107, Rocky Mountain Forest and Experiment Station, Forest Service, U.S. Department of Agriculture, Fort Collins, Colorado.

Stanley, T. D. (2001), 'Wheat from Chaff: Meta-Analysis as Quantitative Literature Review', *The Journal of Economic Perspectives* **15** (Summer), 131–150.

Van den Bergh, J. C. J. M., K. J. Button, P. Nijkamp and G. C. Pepping (1997), *Meta-Analysis in Environmental Economics*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Walsh, R. G., D. M. Johnson and J. R. McKean (1990), 'Nonmarket Values from Two Decades of Research On Recreation Demand', in A. Link and V. K. Smith, eds., *Advances in Applied Micro-Economics*, Vol. 5. Greenwich, CT: JAI Press, pp. 167–193.

Weizman, M. (1998), 'The Noah's Ark Problem', *Econometrica* **66** (November), 1279–1299.

Willig, R. D. (1976), 'Consumer's Surplus Without Apology', *American Economic Review* **66** (September), 589–597.

Willig, R. D. (1978), 'Incremental Consumer's Surplus and Hedonic Price Adjustment', *Journal of Economic Theory* **17**(2), 227–253.

Woodward, R. T. and Y-S. Wui (2001), 'The Economic Value of Wetland Services: A Meta-Analysis', *Ecological Economics* **37**, 257–270.