Modeling relationships between use and nonuse values for surface water quality: A meta-analysis

Robert J. Johnston
Department of Agricultural and Resource Economics, University of Connecticut, Avery Point, Groton, Connecticut, USA

Elena Y. Besedin and Ryan F. Wardwell
Abt Associates, Inc., Cambridge, Massachusetts, USA

Received 3 September 2003; revised 23 October 2003; accepted 31 October 2003; published 19 December 2003.

[1] This paper describes a meta-analysis conducted to estimate relationships between the nonuse components of WTP for surface water quality improvements and a combination of resource, context, and study design attributes, where these attributes include estimated use values for identical improvements. Results are contrasted across four different statistical approaches: ordinary least squares, weighted least squares, multilevel models, and simultaneous equation models. Findings illustrate robust, systematic relationships between a variety of attributes and estimated nonuse WTP, in addition to a systematic empirical relationship between use and nonuse values. Results are promising with regard to the ability of meta-analysis to synthesize information regarding nonuse WTP for water quality improvements and to reveal relationships unapparent from individual stated preference models. *INDEX TERMS:* 6304 Policy Sciences: Benefit-cost analysis; 6329 Policy Sciences: Project evaluation; 6399 Policy Sciences: General or miscellaneous; *KEYWORDS:* meta-analysis, water quality, nonuse value, stated preference, willingness to pay

Citation: Johnston, R. J., E. Y. Besedin, and R. F. Wardwell, Modeling relationships between use and nonuse values for surface water quality: A meta-analysis, *Water Resour. Res.*, *39*(12), 1363, doi:10.1029/2003WR002649, 2003.

1. Introduction

[2] It is generally accepted that natural resource improvements, including improvements to water resources, may generate both use and nonuse values. The specific determinants of use and nonuse values, however, may differ. There is no necessary theoretical link between the magnitudes of use and nonuse values, and in most all cases there is an accompanying lack of behavioral trails from which one may validate nonuse value estimates [*Freeman*, 2003]. Nonetheless, it is possible that estimated use values might in some cases serve as a systematic, empirical indicator of nonuse values for identical resource changes. Here we use meta-analysis to explore empirical relationships which may provide guidance in establishing reasonable expectations for nonuse values. Among these relationships is the relationship between estimated use and nonuse values, in this case for identical improvements in surface water quality.

[3] *Glass* [1976, p. 3] characterizes meta-analysis as "the statistical analysis of a large collection of results for individual studies for the purposes of integrating the findings. It connotes a rigorous alternative to the casual, narrative discussion of research studies which typify our attempt to make some sense of the rapidly expanding research literature" (as discussed by *Poe et al.* [2001, p. 138]). Meta-analysis has a long history in fields such as epidemiology and education, with typical applications to sets of studies conducted under controlled conditions with standardized experimental designs [*Bateman and Jones*, 2003; *Glass et al.*, 1981; *Wolf*, 1986]. However, meta analysis techniques have been increasingly explored by economists as a potential basis for nonmarket benefit transfer [*Bateman and Jones*, 2003].

[4] While many authors advise caution in direct policy applications of meta-analysis [e.g., *Poe et al.*, 2001], such methods are increasingly recognized as a means to assess causes of variability in willingness to pay (WTP) estimates or to model empirical relationships among different types of WTP [*List and Gallet*, 2001; *Santos*, 1998]. Its potential to reveal correlations and systematic causal patterns in existing valuation studies is well known [*Bateman and Jones*, 2003; *Rosenberger and Loomis*, 2000b; *Woodward and Wui*, 2001], as is its potential for encouraging consensus on reasonable magnitudes for resource value estimates [*Poe et al.*, 2001].

[5] The use of meta-analysis to identify systematic relationships between use and nonuse values for surface water quality improvements, accounting for the potential impact of other exogenous influences, may provide at least a preliminary means to evaluate the magnitudes of nonuse values. For example, such analysis could provide an initial means to identify studies in which use/nonuse value ratios are far from the norm; further scrutiny might be called for in such cases. Such use of meta-analysis is similar to the approach used by *List and Gallet* [2001] to assess systematic disparities between actual and hypothetical WTP, and the assessment of use/nonuse value ratios of *Brown* [1993].

[6] Results may also shed light upon whether estimated use and nonuse values, as expected, reveal distinct and systematic correlations with economic, resource, and study design attributes. Contrary findings (e.g., a near-perfect

0043-1397/03/2003WR002649$09.00

Table 1. Studies Incorporated in Nonuse WTP Meta-analysis

Source	Year of Survey	Observations in Meta-data	USDA Region	Survey Method	Water Body	WTP Elicitation	Method for Distinguishing Nonuse WTP[a]
Huang et al. [1997]	1995	2	southeast	telephone	estuary	discrete choice	method 5
Whitehead et al. [1995]	1990	1	southeast	telephone	estuary	iterative bidding	method 1
Sanders et al. [1990]	1983	4	Pacific/mountain	mail	rivers	open ended	method 3
Magat et al. [2000]	1997	1	multiple regions	in person	all freshwater	iterative bidding	method 4
Lant and Roberts [1990]	1987	3	northeast	in person	rivers	payment card	method 2
Roberts and Leitch [1997]	1996	1	multiple regions	mail	wetlands	payment card	method 2
Croke et al. [1986]	1986	3	northeast	telephone	rivers	iterative bidding	method 1
Olsen et al. [1991]	1989	1	Pacific/mountain	telephone	rivers	open ended	method 1
Sutherland and Walsh [1985]	1981	1	Pacific/mountain	mail	rivers and lakes	open ended	method 3
Walsh et al. [1978]	1976	1	Pacific/mountain	in person	rivers	iterative bidding	method 4
Cronin [1982]	1973	1	northeast	in person	rivers	open ended	method 1
Desvousges et al. [1983]	1981	1	northeast	in person	rivers	payment card	method 4
Whitehead and Groothuis [1992]	1991	1	southeast	mail	rivers	open ended	method 1
Kaoru [1993]	1989	1	northeast	mail	estuary	open ended	method 3
Welle [1986]	1985	1	northeast	mail	all freshwater	open ended	method 2
Rowe et al. [1985]	1985	2	Pacific/mountain	mail	rivers	open ended/payment card	method 5
Mitchell and Carson [1981]	1980	1	multiple regions	in-person	all freshwater	payment card	method 1
Clonts and Malone [1990]	1987	1	southeast	telephone	rivers	iterative bidding	method 2
Bockstael et al. [1989]	1984	1	northeast	telephone	estuary	discrete choice	method 2
Whittington et al. [1994]	1993	1	southern plains	mail	estuary	discrete choice	method 2

[a]Method 1 is nonuse values estimated as the total WTP for nonusers. Method 2 is nonuse value estimated using responses to a separate nonuse value question in the survey. Method 3 is nonuse values estimated through apportionment of total WTP among categories of value by survey respondents. Method 4 is nonuse values estimated as the total WTP of a survey subsample who were asked to assume they would not use the resource being valued. Method 5 is nonuse values estimated as the total WTP of the sample of users minus estimated WTP for the direct use of the resource estimated based on revealed preference data.

correlation between use and nonuse values, with other attributes revealing little significant or systematic influence; statistical relevance of variables that should have no impact) might call into question the validity of underlying value estimates, and the ability of stated preference surveys to truly distinguish use and nonuse values [*Santos* 1998].

[7] This paper describes the methodology and results of a meta-analysis conducted to estimate relationships between the nonuse components of WTP for surface water quality improvements and a combination of economic, resource, and study design attributes, where these attributes include estimated use values for identical improvements. The primary objectives of this analysis are to systematically contrast the magnitudes of use and nonuse values associated with surface water quality, and to assess other systematic relationships that influence estimated nonuse WTP.

2. Data and Conceptual Approach

[8] As a largely empirical exercise, the applicability of meta-analysis to any particular research question is dependent on the ‘‘number, quality and comparability’’ of studies incorporated in the meta-data [*Bateman and Jones*, 2003, p. 237; *Desvousges et al.*, 1998]. In social science applications, there is typically a tradeoff between the number of studies that may be incorporated in the data, and the degree of heterogeneity in study approaches (and resulting stress on the data) which the researcher is willing to accept [*Bateman and Jones*, 2003]. Here we restrict the universe of available studies using four factors: (1) a requirement that the study allows one to separate or otherwise distinguish use and nonuse values for surface water improvements; (2) a requirement that surface water quality improvements are related to improvements in aquatic life and/or habitat in a water body that provides recreational fishing; (3) a limitation to studies conducted in the U.S.; and (4) a limitation to studies conducted using approaches generally accepted in the literature.

[9] The resulting meta-data comprise 29 observations gathered from 20 unique studies published between 1978 and 2000. Table 1 summarizes the characteristics of each study. The number of observations exceeds the number of studies because five studies provide more than one estimate of both use and nonuse value. Multiple estimates of nonuse and use value from a single study are available due to in-study variations in the quantity of the amenity provided [*Sanders et al.*, 1990], the degree of quality change [*Croke et al.*, 1986; *Lant and Roberts*, 1990], and the techniques employed to estimate nonuse value [*Huang et al.*, 1997; *Rowe et al.*, 1985]. While the number of observations is small compared to prior meta-analyses [e.g., *Dalhuisen et al.*, 2003; *Poe et al.*, 2001; *Rosenberger and Loomis*, 2000a, 2000b; *Santos*, 1998], we follow *Bateman and Jones* [2003] and favor the reduction of stress on the data over expansion of the meta-data.

[10] The isolation of nonuse and use values associated with water resources draws on the general approach of *Brown* [1993]. On the basis of the characteristics of the original studies, nonuse values were isolated using five methods: (1) nonuse values identified as the total WTP for nonusers; (2) responses to a separate nonuse value question in the survey; (3) apportionment of total WTP among categories of value by survey respondents; (4) nonuse values estimated as the total WTP of a survey subsample who were asked to assume they would not use the resource being valued; and (5) total WTP of the sample of users minus estimated WTP for the direct use of the resource estimated based on revealed preference data. Corresponding use values were obtained using analogous methods. In most cases, results of these methods were

supplied by the original studies and appropriate value estimates were simply transcribed from the published reports. Table 1 illustrates the method applied to each study in the data.

[11] As shown in Table 1, method 1 is among the most common methods used to isolate nonuse values. This method estimates total WTP for user and nonuser subsamples of respondents. Despite variations across studies, respondents were generally characterized as nonusers of the resource if the respondent did not use the resource during some defined historical period, or did not expect to use the resource during some defined future period. We note that this method of distinguishing use and nonuse values may underestimate nonuse values among resource users, as it implies that nonuse values are identical for users and nonusers [*Whitehead and Blomquist*, 1991]. To account for this possibility, the subsequent analysis includes a variable identifying nonuse estimates calculated using this method.

[12] Studies using method 2 ask respondents to report separate WTP bids for distinct components of total value. Respondents typically answer a set of valuation questions designed to estimate value for separate components of total value (e.g., use, nonuse). For example, *Lant and Roberts* [1990] obtain bids from each respondent for both recreational use value and existence value. Method 3 requires that respondents first report their total WTP and then allocate this value among different categories. For example, *Rowe et al.* [1985] ask respondents to allocate WTP among use, existence, and bequest values.

[13] Method 4 calculates nonuse WTP by eliciting resource values under conditions in which respondents are asked to assume they would not use the resource being valued. For example, *Walsh et al.* [1978] ask, "If it were certain you would not use the South Platte River Basin for water-based recreation, would you be willing to add ____ cents on the dollar to present sales taxes every year, just to know clean water exists at level A as a natural habitat for plants, fish, wildlife, etc.?" Method 5 combines revealed and stated preference data to estimate use and nonuse value. For example, *Huang et al.* [1997] combine stated preference surveys with travel cost methods to estimate both total and use WTP; nonuse value is calculated as the difference between these two estimates.

[14] The dependent variable in all estimated models is estimated nonuse value for water quality improvement. On the basis of economic theory and findings from the literature, we expect that various factors may influence nonuse values. Following *Poe et al.* [2001], we assume that these may include study design effects (e.g., WTP elicitation methods, survey format), economic variables (e.g., income), and resource characteristics (e.g., type of water body, degree of quality change, geographical scale of improvements). All models also include estimated use values for identical water quality improvements as a rhs variable or indicator of nonuse WTP. Table 2 characterizes independent variables included in one or more estimated models.

[15] Although the interpretation and calculation of most independent variables (Table 2) are relatively transparent, there are a few variables for which additional explanation is required. Primary among these are variables characterizing surface water quality and its measurement (*wq_ladder*, *wq_change*). The majority of observations in the meta-data characterize quality changes using variants of the Resources for the Future (RFF) water quality ladder [e.g., see *Mitchell and Carson*, 1989]. However, twelve observations provide water quality measures using other, primarily descriptive, means. To allow inter-observation comparisons of water quality change using a single scale, it was necessary to map the verbal descriptions of water quality present in these nonladder studies onto the standardized water quality ladder. In most cases the descriptions of water quality (present in the studies that did not apply the water quality ladder) rendered this procedure straightforward. The variable *wq_change* characterizes the difference between baseline and post-policy water quality, measured in water quality ladder units. To distinguish those studies in which these water quality ladder measurements were an original component of the survey instrument, we introduce the binary variable *wq_ladder* (Table 2).

[16] The proxy variable *income* also requires explanation. While some studies in the meta-data report the average income of survey respondents, many do not provide this information. For studies that do not report mean household income, income is approximated using median household income as reported in the U.S. Census, calculated for the geographic area(s) that most closely approximate that of the study area, for the Census year closest to the year in which the study was conducted. All data are adjusted to 2002 dollars. While this approximation is necessary in order to allow a measure of income to be included in the model, it should be emphasized that the resulting instrument for respondent income represents at best a rough approximation; results with respect to this variable should be interpreted accordingly.

3. Empirical Methods

[17] Past meta-analyses have incorporated a range of different statistical methods, with none universally accepted as superior [e.g., *Poole and Greenland*, 1999; *Bateman and Jones*, 2003]. Standard linear regression, multilevel models (i.e., random-effects or random-coefficients models), and weighted regression models are common. (Within much of the meta-analysis literature, explicit weighting of observations to account for research quality is also common [*Wolf*, 1986; *Santos*, 1998]. However, for meta-analyses involving WTP estimates or resource values, clear and usable measures of research quality are often difficult to obtain. For example, standard errors of WTP estimates are not commonly reported, particularly in older studies. For this reason, studies in the environmental economics literature do not often provide explicit weights with regard to measures of research quality.) Given the variety of statistical approaches present in the literature and the lack of consensus regarding the most appropriate approach(es), we illustrate estimated results using four statistical methods: (1) ordinary least squares, (2) weighted least squares, (3) multilevel (random-effects) models, and (4) simultaneous equation models (two-stage least squares).

[18] The ordinary least squares (OLS) model represents the simplest approach to the data. As observations are clustered by study author (i.e., numerous studies provide more than one observation in the data) and estimation is based on a pooled sample with mean WTP as the

Table 2. Meta-analysis Variables and Descriptive Statistics

Variable	Description	Units and Measurement	Mean (SD)
ln_nonuse	natural log of nonuse values for specified resource improvements	natural log of dollars (range: 0.74 to 5.49)	4.02 (0.99)
use_value	estimated use values for specified resource improvements	dollars (range: $5.17 to $275.09)	63.94 (67.49)
wq_change[a]	change in mean water quality, specified on the RFF water quality ladder [*Mitchell and Carson*, 1989]; see text for additional information	water quality ladder units (range: 0.5 to 5.75)	2.41 (1.39)
base_wq[a]	baseline water quality, specified on the RFF water quality ladder.	water quality ladder units (range: 0 to 7)	4.18 (2.49)
estuary[a]	binary (dummy) variable indicating that resource change takes place in an estuary	binary (range: 0 or 1)	0.21 (0.41)
wetland[a]	binary (dummy) variable indicating that resource change takes place in a water body with wetlands as a primary stated component (e.g., a combined wetland/reservoir complex)	binary (range: 0 or 1)	0.03 (0.19)
rivers_lakes[a]	binary (dummy) variable indicating that resource changes take place in rivers and/or lakes	binary (range: 0 or 1)	0.66 (0.48)
saltponds[a]	binary (dummy) variable indicating that resource changes take place in salt ponds	binary (range: 0 or 1)	0.03 (0.19)
num_riv_lake[a]	number of rivers and/or lakes affected by policy; if unspecified *num_riv_lake* = 0; equivalent to a multiplicative interaction between *rivers_lakes* and the number of water bodies affected (see text for additional information)	number of water bodies (range: 0 to 15)	2.07 (4.29)
nocount[a]	Binary (dummy) variable indicating that *number* = 0	binary (range: 0 or 1)	0.31 (0.47)
sm_nocount[a]	multiplicative interaction between *no_count* and a binary (dummy) variable identifying studies that specify water quality changes over smaller local or regional scales compared to the default of larger (state or nationwide) changes	binary (range: 0 or 1)	0.24 (0.44)
income[a]	mean income of survey respondents (see additional comments in text)	dollars (range: 30396 to 137693)	48267.76 (18881.67)
year_indx[a]	year in which the study was conducted, converted to an index by subtracting 1970	year index (range: 3 to 27)	16.31 (5.64)
discrete[a]	binary (dummy) variable indicating that WTP was estimated using a discrete choice survey instrument	binary (range: 0 or 1)	0.14 (0.35)
mail	binary (dummy) variable indicating that WTP was estimated using a mail survey instrument	binary (range: 0 or 1)	0.41 (0.50)
phone	binary (dummy) variable indicating that WTP was estimated using a telephone survey instrument	binary (range: 0 or 1)	0.31 (0.47)
wq_ladder[a]	binary (dummy) variable indicating that the original survey reported resource changes using a water quality ladder	binary (range: 0 or 1)	0.59 (0.50)
northeast[a]	binary (dummy) variable indicating that survey was conducted in the USDA northeast region	binary (range: 0 or 1)	0.38 (0.49)
southeast[a]	binary (dummy) variable indicating that survey was conducted in the USDA southeast region	binary (range: 0 or 1)	0.17 (0.38)
pacif_mount[a]	binary (dummy) variable indicating that survey was conducted in the USDA Pacific/mountain region	binary (range: 0 or 1)	0.31 (0.47)
south_plain[a]	binary (dummy) variable indicating that survey was conducted in the USDA southern plains region	binary (range: 0 or 1)	0.03 (0.19)
nu_meth	binary (dummy) variable indicating that nonuse WTP was estimated using method 1 (see text)	binary (range: 0 or 1)	0.48 (0.51)
fish_chg[b]	binary (dummy) variable identifying studies in which the survey quantifies specific changes in recreational fishing uses	binary (range: 0 or 1)	0.31 (0.47)
other_chg[b]	binary (dummy) variable identifying studies in which the survey quantifies specific changes in nonfishing uses	binary (range: 0 or 1)	0.76 (0.44)
revealed[b]	binary (dummy) variable identifying studies using revealed preference methods to estimate use values	binary (range: 0 or 1)	0.14 (0.35)

[a]Variable used as an instrument for use value in the 2SLS model and also included as a rhs variable in at least one nonuse value equation.
[b]Variable only used as an instrument for use value in the 2SLS model.

dependent variable, we follow *Smith and Osborne* [1996] and *Dalhuisen et al.* [2003] in using a Huber-White robust covariance estimator to estimate standard errors. As noted by *Smith and Osborne* [1996, p. 293], "this approach treats each study as the equivalent of a sample cluster with the potential for heteroskedasticity. . . .across clusters."

[19] The weighted least squares (WLS) approach follows that of *Poe et al.* [2001]. (This approach represents a standard application of WLS, in which the standard OLS estimator $(X'X)^{-1}X'Y$ is replaced by the WLS estimator $(X'WX)^{-1}X'WY$, where W is a diagonal matrix comprised of weights for each observation.) Weights are defined such that each unique study is given identical weight (i.e., weights on multiple observations within each study sum to one; for studies with only one observation, the weight is equal to one). This ensures that a study providing multiple observations within the meta-data does not receive undue weight in the regression analysis. Again following *Poe et al.* [2001], standard errors are calculated using the Huber-White robust covariance estimator.

Table 3. Estimated Meta-analyses of Nonuse WTP for Surface Water Quality—*Ln*(*Non_Use*)[a]

	Model One: OLS (Unrestricted)	Model Two: OLS (Restricted)	Model Three: WLS	Model Four: Multilevel	Model Five: 2SLS
intercept	4.8740[b] (2.5749)	4.5213[c] (0.6238)	4.4010[c] (0.6568)	4.5213[c] (0.3981)	4.6809[c] (1.1136)
use_value	0.0124 (0.0206)	0.0065[c] (0.0017)	0.0065[c] (0.0015)	0.0065[c] (0.0011)	0.0058[d] (0.0019)
wq_change	0.1459 (0.1062)	0.1384[c] (0.0377)	0.1342[d] (0.0561)	0.1384[c] (0.0241)	0.1430 (0.0966)
base_wq	−0.1248 (0.0987)	−0.0418 (0.0255)	−0.0312 (0.0265)	−0.0418[b] (0.0163)	−0.0387 (0.0507)
estuary	−1.8267 (6.5244)	1.0389[c] (0.2514)	1.1048[c] (0.2777)	1.0389[c] (0.1604)	1.0212 (0.7937)
saltpond	−4.3426 (12.9427)	0.7102 (0.5315)	0.6662 (0.5349)	0.7102[b] (0.3392)	0.8221 (0.8967)
wetland	−2.5889 (3.9446)	−2.2164[c] (0.3601)	−2.2256[c] (0.3679)	−2.2164[c] (0.2298)	−2.1480[d] (0.7856)
rivers_lakes	−2.3448 (4.9631)	0.0895 (0.2013)	0.1729 (0.2509)	0.0895 (0.1284)	0.0746 (0.6361)
num_riv_lake	0.0188 (0.0800)	0.0316[b] (0.0163)	0.0258[b] (0.0125)	0.0316[d] (0.0104)	0.0296 (0.0274)
nocount	−0.2260 (5.3877)	1.2903[c] (0.3697)	1.3154[c] (0.3527)	1.2903[c] (0.2359)	1.3518[b] (0.7069)
sm_nocount	0.6140 (3.4263)	−0.8769[c] (0.2562)	−0.9840[c] (0.2885)	−0.8769[c] (0.1635)	−0.8590 (0.7148)
northeast	3.2301 (4.0634)	1.1700[c] (0.2797)	1.1768[c] (0.2530)	1.1700[c] (0.1785)	1.1094[d] (0.4079)
southeast	2.1765 (2.8870)	0.1363 (0.2223)	0.1457 (0.2081)	0.1363 (0.1418)	0.1838 (0.3820)
pacif_mount	1.5304 (1.6835)	–	–	–	–
south_plain	3.8270 (8.7134)	–	–	–	–
income	0.00002 (0.00006)	–	–	–	–
year_indx	−0.0735[b] (0.0419)	−0.0562[b] (0.0319)	−0.0532 (0.0313)	−0.0562[d] (0.0204)	−0.0655 (0.0412)
discrete_ch	−1.5335 (3.7375)	−0.8380[d] (0.3306)	−0.9318[d] (0.3345)	−0.8380[d] (0.2109)	−0.7403 (0.6196)
mail	0.1172 (2.0052)	–	–	–	–
phone	0.3838 (3.0194)	–	–	–	–
wq_ladder	−2.0167 (1.4463)	−1.2828[c] (0.1766)	−1.1986[c] (0.2597)	−1.2828[c] (0.1127)	−1.2796[d] (0.4393)
nu_meth	−0.9312 (2.1922)	−0.4055[c] (0.0944)	−0.4713[c] (0.1188)	−0.4055[c] (0.0602)	−0.4056[b] (0.2084)
R^2	0.9347	0.9209	0.9525	–	0.9199
N	29	29	29	29	29
Wald test of restrictions (df = 5,19)	–	0.73 (p = 0.61)	–	–	–
σ^2_u	–	–	–	1.51×10^{-18}	–
χ^2 for $\sigma^2_u = 0$	–	–	–	0.00 (p = 1.00)	–
χ^2 for Hausman test[e]					1.08 (p = 0.30)

[a]For each variable, numbers outside parenthesis are coefficients, and numbers inside parenthesis are standard errors.
[b]$p < 0.10$.
[c]$p < 0.01$.
[d]$p < 0.05$.
[e]Hausman test specified to assess the null hypothesis of OLS consistency relative to 2SLS.

[20] The multilevel model follows the general approach of *Bateman and Jones* [2003]. Multilevel (or hierarchical) models are also variously denoted random-effects or random-coefficients models, and are described in detail elsewhere [*Goldstein*, 1995; *Singer*, 1998]. The fundamental distinction between these and classical linear models is the two-part modeling of the equation error to account for hierarchical data. Here, as the meta-data are comprised of (in some cases) multiple observations per study, there is a possibility of correlated errors among observations that share a common study or author.

[21] A common approach to modeling such potential correlation is to divide the residual variance of estimates into two parts, a random error that is independently and identically distributed across all studies and for each observation, and a random effect that represents systematic variation related to each study. The model is estimated as a two-level hierarchy, with level one corresponding to nonuse value estimates (individual observations), and level two corresponding to individual studies. The random-effect may be interpreted as a deviation from the mean equation intercept associated with individual studies [*Bateman and Jones*, 2003]. Here the model is estimated using maximum likelihood assuming that random effects are distributed multivariate normal. (A fixed effects model is precluded here due to the small sample size.) Following *Bateman and Jones* [2003], who argue against weighting observations in multilevel meta-analysis, observations remain unweighted. As in the OLS model above, covariances are obtained using the Huber-White covariance estimator.

[22] The third statistical approach explicitly recognizes that use value (a rhs variable in the nonuse value regression model) is measured with error, and may be most appropriately treated as endogenous. That is, estimated use value is stochastic and may be correlated with the equation error. In such cases classical least squares regression will produce biased and inconsistent parameter estimates [*Greene*, 2003]. Accordingly, the third and final statistical approach uses instrumental variables (two-stage least squares, or 2SLS) to estimate the model. The estimated equation models nonuse value as a function of use value (the endogenous, instrumented variable) and a set of additional exogenous variables. Observations are again unweighted.

[23] The contrast of these four statistical models (OLS,WLS, multilevel model, 2SLS) allows assessments of the robustness of model results. That is, the application of distinct modeling approaches may help identify those effects that are highly sensitive to changes in model specification and estimation method, versus those that are robust to such changes. It also allows assessment of which of the four approaches is most appropriate, given the existing meta-data.

4. Model Results

[24] Results are shown in Table 3. Five models are illustrated. Model one is an unrestricted OLS model. Model two is a restricted OLS model, identical to the large model save for the exclusion of five independent variables. These excluded variables include dummy variables distinguishing two geographic regions (*pac_mount*, *south_plain*; Table 2),

surveys implemented using mail and phone methods (*mail*, *phone*; Table 2) and the proxy for mean respondent income (*income*, Table 2). (While one might choose to retain the variables *income*, *mail*, and *phone* based on prior arguments in the literature regarding the potential relevance of such factors [e.g., *Arrow et al.*, 1993], we chose to exclude these variables due to their clear lack of explanatory power and small number of degrees of freedom in the unrestricted model.) A Wald test of the joint statistical significant of these restrictions fails to reject the null hypothesis of zero joint influence (F = 0.73; p = 0.61). Model three is the WLS model, with variables identical to those in the restricted OLS model. Models four and five represent the multilevel and 2SLS models; independent variables are again identical to those of model two. Instruments for the 2SLS model are shown in Table 2; only the nonuse equation is formally estimated, eliminating the necessity of specifying a functional form for the remaining (use value) equation [*StataCorp*, 2001]. The nonuse equation satisfies both the rank and order conditions for identification. (Instruments for use value that are excluded from the nonuse equation include *fish_chg* (a binary variable identifying studies in which the survey quantifies specific changes in recreational fishing), *other_chg* (a binary variable identifying studies in which the survey quantifies specific changes in nonfishing uses), and *revealed* (a binary variable identifying studies in which use values are estimated using revealed preference methods).)

[25] In all models, the dependent variable is the natural log of estimated nonuse WTP. All rhs variables are linear, resulting in a semi-log functional form which reflects the underlying relationship

$$V_i = e^{\beta_0 + \sum_{j=1}^{J} \beta_j X_{ij} + \varepsilon_i} \tag{1}$$

where V_i represents estimated nonuse WTP for observation i, the X_{ij} are rhs (explanatory) variables, the β_j are associated parameters to be estimated, and ε_i represents the equation error. (For the multilevel model, the specification in effect replaces ε_i with $\varepsilon_i = \mu_k + u_{ik}$, where both μ_k and u_{ik} are assumed normally distributed with mean zero. The u_{ik} are treated as errors that are iid across all observations. The μ_k may be interpreted as study-level effects, where k subscripts individual studies. These effects are assumed iid across different studies but may be correlated within each study [cf. *Bateman and Jones*, 2003].) Given equation (1), the expected percentage change in nonuse WTP resulting from a single unit change in the jth variable is given by $(e^{\beta j} - 1) \times 100$, where $(e^{\beta j} - 1)$ approaches β_j as β_j approaches zero [*Chicoine*, 1981].

4.1. Restricted Model Specification Tests

[26] Although results for all models are illustrated, specification tests suggest that the restricted OLS specification may be best suited to the data. A likelihood ratio test of the multilevel, random effects model (strongly) fails to reject homogeneity in errors across studies ($\chi^2 = 0.00$, p = 1.00), indicating a lack of significant improvement over the equal-effects OLS model. (An analogous test was conducted for models estimated without robust covariance estimation methods; results were equivalent.) This result is not particularly surprising, given similar findings (i.e., a lack of statistically significant panel effects) of *Rosenberger and Loomis* [2000a] and *Bateman and Jones* [2003]. (*Rosenberger and Loomis* [2000a] provide a more extensive discussion of potential tests of panel effects in meta-data.) A Hausman test of the 2SLS model generates similar results, failing to reject the null hypothesis that OLS is a consistent estimator of the nonuse value equation ($\chi^2 = 1.08$, p = 0.30). On the basis of these results we conclude that neither random-effects nor 2SLS offer statistically significant improvements over the simpler OLS model.

[27] The choice between WLS and OLS is most appropriately made on theoretical grounds. However, there is no clear consensus in the literature regarding the choice of weighted versus unweighted meta-models. While weighting of observations in meta-analysis is common [e.g., *Mrozek and Taylor*, 2002; *Poe et al.*, 2001], *Bateman and Jones* [2003] argue that such methods are inappropriate. Here, the choice of WLS versus OLS is largely symbolic, as results from the two models are similar. Hence, given the lack of an unambiguous justification for applying WLS, we emphasize results of the simpler OLS specification.

[28] Statistical test results notwithstanding, the four restricted models generate similar results. All models are significant at $p < 0.01$, and despite the small sample size R^2 values (>0.90 in all cases) compare favorably to those of prior meta-analyses [e.g., *Bateman and Jones*, 2003; *Poe et al.*, 2001; *Santos*, 1998]. Signs of significant parameter estimates correspond with intuition. Perhaps the most notable difference among models is the relatively larger standard errors evident in the 2SLS model, compared to other restricted specifications.

4.2. Determinants of Nonuse WTP for Surface Water Quality: Use Values and Resource Attributes

[29] Despite variations in methods and nonuse WTP reported by individual studies, we find a strong systematic element of nonuse WTP variation. Findings illustrate systematic and statistically significant relationships between a variety of resource, context, and scale attributes and reported nonuse WTP. With few exceptions, the signs and magnitudes of parameter estimates are similar across all restricted model specifications (Table 3). However, based on the results of statistical tests noted above, the following discussion is based on findings of the restricted OLS specification.

[30] As expected, increases in use values are associated with increases in estimated nonuse values ($p < 0.01$). The magnitude of OLS parameter estimates indicates that each \$1 increase in use values is associated with a 0.67% increase in nonuse values. At a mean estimated nonuse value of \$75.63, this implies a \$0.51 increase in nonuse values for every \$1 increase in use values, holding all else constant. This finding corresponds with the more general conclusion of *Brown* [1993, p. 173] that use of a resource appears to "enhance. . .nonuse values."

[31] Measures of resource quality also influence nonuse WTP. The variable *wq_change* represents the change in surface water quality in RFF ladder units. The model reveals a statistically significant ($p < 0.01$) and positive relationship between *wq_change* and nonuse WTP, indicating that larger water quality improvements are associated with higher nonuse WTP estimates. The variable *base_wq* represents the baseline water quality from which change would occur. Although the associated parameter estimate is of the expected negative sign (indicating that nonuse WTP for

water quality improvements diminishes as baseline quality increases), it is only significant at $p < 0.12$.

[32] Relationships between nonuse WTP and the type of water body under consideration are less clear, with only two of four parameter estimates statistically significant. Among significant effects, the hypothesis that nonuse values for estuarine water quality changes are the same as those for (the default of) unspecified freshwater bodies (*estuary* = 1) may be rejected at $p < 0.01$. The same null hypothesis for wetlands (*wetlands* = 1) may be rejected at $p < 0.01$. In contrast, WTP for water quality improvements in salt ponds, rivers, and lakes cannot be distinguished from the default category at $p < 0.10$.

[33] Model results further suggest that the geographic and numerical scale of resource improvements influences nonuse WTP. Studies in which water quality changes are specified over geographic regions (e.g., states, counties, or watersheds) with no explicit water body count are identified by the binary variable *nocount*, to distinguish these studies from others in which specific numbers of water bodies are indicated by the survey. A second binary variable, *sm_nocount*, further distinguishes studies in which both (1) the number of water bodies is unspecified and (2) the region affected by the policy is less than statewide. Both variables are statistically significant. Results indicate that studies in which the number of affected water bodies was unspecified (*nocount* = 1) are associated with higher nonuse WTP. This positive effect is reduced (but not eliminated) for regions of smaller size (*sm_nocount* = 1).

[34] Only a subset of studies (surveys) indicated that a specific number of water bodies would be affected by policy actions. Among these studies, only those addressing rivers and lakes (*rivers_lakes* = 1) provide variation in the number of identified water bodies considered by respondents. The variable *num_riv_lake* is a multiplicative interaction between the binary variable *rivers_lakes* and the number of water bodies under consideration (Table 2). The associated parameter estimate is statistically significant and indicates that nonuse WTP increases with the number of water bodies considered. This result, combined with the statistical significance of *wq_change* noted above, suggests that that WTP values (in this case nonuse values for water quality improvements) are sensitive to scope, both in terms of the number of water bodies considered and the magnitude of water quality change. This result coincides with prior meta-analytic findings of *Smith and Osborne* [1996] concerning WTP for visibility in national parks.

[35] Finally, the region in which improvements occur influences nonuse WTP. However, statistically significant impacts ($p < 0.01$) are only associated with the regional variable *northeast* (Table 2). For studies characterized by *northeast* = 1 (38% of observations), model results show a statistically significant ($p < 0.01$) and substantial increase in nonuse WTP. Statistically significant effects could not be established for other U.S. regions. The remaining regional dummy variables are either excluded from the restricted model due to lack of statistical significance (see joint tests above), or are insignificant in the restricted model.

4.3. Determinants of Nonuse WTP for Surface Water Quality: Survey Design Characteristics

[36] A small number of survey design characteristics can be shown to influence nonuse WTP. Among these is the year in which a study was conducted, with later studies associated with reduced nonuse WTP estimates. This is the expected result, as the focus on ongoing improvements in survey design has often been on the reduction of survey biases that would otherwise result in an overstatement of true WTP [*Arrow et al.*, 1993]. Model results also reveal a statistically significant ($p < 0.05$) decrease in nonuse WTP associated with discrete choice instruments. This finding is of note given contrary findings of *Boyle et al.* [1996] and *Ready et al.* [1996]. Hence, while some prior research has shown that discrete choice methods may be associated with higher WTP estimates, results here suggest that this pattern may not be universal.

[37] Models results indicate a statistically significant ($p < 0.01$) and negative effect associated with the explicit use of a water quality ladder within the survey to represent water quality change (*wq_ladder*). This effect suggests that studies using the water quality ladder may generate lower nonuse WTP estimates, perhaps due to the capacity of such scales to clarify the specific magnitude and implications of water quality change, and hence (perhaps) reduce methodological misspecification or symbolic biases that might act to systematically inflate estimated WTP.

[38] Finally, the method of distinguishing use and nonuse values (*nu_method*, a binary variable identifying studies that estimate nonuse values using method 1: surveys administered to user and nonuser populations) appears to influence nonuse WTP. Studies applying distinct user and nonuser surveys to estimate nonuse values generated systematically lower estimates of nonuse WTP. This is the expected result, given *Whitehead and Blomquist*'s [1991] argument that such methods may underestimate nonuse values of resource users.

[39] Although certain elements of survey design and research methodology appear to influence nonuse WTP, the survey administration method (e.g., mail, phone) appears to have no systematic influence in our meta-data. These variables are dropped from the unrestricted model due to a clear lack of statistical significance. This finding is notable given the advice of the NOAA Blue Ribbon Panel [*Arrow et al.*, 1993] that "face-to-face interviews are usually preferable." Findings here fail to find any influence on nonuse WTP associated with the use of mail or telephone surveys, contrasted to the default category of an in-person survey.

4.4. Implications and Discussion

[40] Despite the attention provided to the methodological details of meta-analysis [e.g., *Poole and Greenland*, 1999; *Bateman and Jones*, 2003], most findings here are robust to changes in statistical approach (Table 3). Indeed, hypothesis tests indicate that none of the models provide significant improvements over OLS estimated with robust standard errors. This is not to suggest that model development is unimportant, but rather the more hopeful contention that a wide range of statistical approaches to meta-analysis may provide valid and useful findings.

[41] Results of the meta-analysis reveal a number of insights regarding nonuse WTP for water quality improvements. Among the most fundamental of these is that nonuse values (for surface water quality improvements) appear to vary systematically with a range of exogenous variables characterizing resource quality, scope, and context. Of more specific significance, the model finds a clear empirical

relationship between use and nonuse values. This relationship is purely empirical; it is not required by neoclassical theory. Nonetheless, findings suggest that use values may in some cases serve as a statistically significant, systematic indicator of nonuse values. Such findings, if found to hold in additional and larger-scale research, may allow estimated use values for resource improvements to provide a preliminary yardstick by which policymakers may assess the magnitudes of nonuse values. Such methods may help researchers identify outlier nonuse WTP estimates that warrant additional scrutiny prior to their use to inform policy. Conversely, such results may also help researchers establish cases in which nonuse WTP estimates are within a reasonable range.

[42] Despite the strong relationship between use and nonuse WTP, the positive and significant intercept coefficient, together with statistically significant parameter estimates associated with other exogenous variables, suggests that nonuse WTP may be positive even when use values are negligible. This result indicates that estimated use and nonuse WTP are not exactly proportional across studies, and adds verisimilitude to the argument that stated preference methods can estimate meaningful differences between use and nonuse values. It further suggests that these two measures represent unique and distinguishable elements of total value.

[43] An additional model finding of potential relevance is that nonuse values appear sensitive to a relatively small number of survey design characteristics, and that such effects do not dominate systematic components of nonuse WTP. While more extreme critiques of stated preference methods may characterize estimated values as being almost entirely determined by features of survey design, the illustrated models find that many aspects of survey design have ambiguous or negligible impacts on stated WTP estimates. In contrast, many measures of resource quality and context are shown to have statistically significant and systematic impacts.

5. Conclusion

[44] The literature is clear regarding the potential pitfalls in the direct use of meta-analyses for benefit transfer [*Desvousges et al.*, 1998; *Poe et al.*, 2001]. Nonetheless, the systematic component of nonuse WTP found here is promising with regard to the use of meta-analysis to guide in preliminary evaluations of resource value estimates. While additional work is needed prior to the application of such results to assess nonuse value magnitudes in policy contexts, model results offer insights regarding systematic components of nonuse WTP for water quality improvements. These insights provide encouraging evidence with regard to the use of stated preference methods to estimate nonuse WTP (e.g., sensitivity of nonuse values to scope).

[45] Of particular focus here is the relationship between use and nonuse values for surface water quality improvements. Findings indicate that nonuse values are strongly correlated with use values. However, nonuse values also vary systematically with a variables characterizing resource quality, scope, and context. Hence, while use values may provide a preliminary indicator by which nonuse values may be assessed, there are additional systematic components of nonuse values that are independent of use values.

[46] Model findings are of course relative to the specific case studies considered. Additional research will be required to assess whether similar findings apply to nonwater resources, or water resources in a broader context. This is particularly germane in the present case, given the relatively small sample size of the meta-data. All statistical findings must be viewed within the context of a small-sample dataset, with all the appropriate caveats. Nonetheless, the systematic components of WTP found in this and other studies suggests that meta-analyses may offer a promising avenue of research for those seeking to evaluate stated preference WTP estimates. More specifically, model results here suggest that meta-analysis can provide a useful means to synthesize information regarding systematic determinants of nonuse WTP.

[47] **Acknowledgments.** This research was partially funded under U.S. EPA contract 68-C-99-239. Although partially prepared with EPA funding, this article has not been reviewed by U.S. EPA. Opinions belong solely to the authors and do not necessarily reflect the views or policies of U.S. EPA or imply endorsement by the funding agency. Address correspondence to Robert J. Johnston at the Connecticut Sea Grant College Program, University of Connecticut at Avery Point, 1080 Shennecossett Road, Groton, CT 06340-6048 or rjohnston@canr.uconn.edu.

References

Arrow, K., R. Solow, E. Leamer, P. Portney, R. Rander, and H. Schuman, Report of the NOAA Panel on Contingent Valuation, *Fed. Regist.*, *58*, 4602–4614, 1993.

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 Econ.*, *79*(2), 235–258, 2003.

Bockstael, N. E., K. E. McConnell, and I. E. Strand, Measuring the benefits of improvements in water quality: The Chesapeake Bay, *Mar. Resour. Econ.*, *6*(1), 1–18, 1989.

Boyle, K. J., F. R. Johnson, D. W. McCollum, W. H. Desvousges, R. H. Dunford, and S. P. Hudson, Valuing public goods: Dicrete versus continuous contingent-valuation responses, *Land Econ.*, *72*(3), 381–396, 1996.

Brown, T. C., Measuring nonuse value: A comparison of recent contingent valuation studies, *W-133 Sixth Interim Rep.*, Dep. of Agric. and Appl. Econ., Univ. of Ga., Athens, 1993.

Chicoine, D. L., Farmland values at the urban fringe: An analysis of sales prices, *Land Econ.*, *57*(3), 353–362, 1981.

Clonts, H. A., and J. W. Malone, Preservation attitudes and consumer surplus in free flowing rivers, in *Social Science and Natural Resource Recreation Management*, edited by J. Vining, pp. 301–317, Westview, Boulder, Colo., 1990.

Croke, K., R. G. Fabian, and G. Brenniman, Estimating the value of improved water quality in an urban river system, *J. Environ. Syst.*, *16*(1), 13–24, 1986.

Cronin, F. J., Valuing nonmarket goods through contingent markets, *Rep. PNL 4255*, Pac. Northwest Lab., Richland, Wash., 1982.

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 Econ.*, *79*(2), 292–308, 2003.

Desvousges, W. H., V. K. Smith, and M. P. McGivney, A comparison of alternative approaches for estimating recreation and related benefits of water quality improvements, report, Econ. Anal. Div., U.S. Environ. Prot. Agency, Washington, D. C., 1983.

Desvousges, W. H., F. R. Johnson, and H. S. Banzhaf, *Environmental Policy Analysis With Limited Information: Principles and Applications of the Transfer Method*, Edward Elgar, Northampton, Mass., 1998.

Freeman, A. M., III, The measurement of environmental and resource values: Theory and methods, report, Resour. for the Future, Washington, D. C., 2003.

Glass, G. V., Primary, secondary, and meta-analysis of research, *Educ. Res.*, *5*(10), 3–8, 1976.

Glass, G. V., B. Shaw, and M. L. Smith, *Meta-Analysis in Social Research*, Sage, Beverly Hills, Calif., 1981.

Goldstein, H., *Multilevel Statistical Models*, 2nd ed., Edward Arnold, London, 1995.

Greene, W. H., *Econometric Analysis*, 5th ed., Prentice-Hall, Old Tappan, N. J., 2003.

Huang, J. C., T. C. Haab, and J. C. Whitehead, Willingness to pay for quality improvements: Should revealed and stated preference data be combined?, *J. Environ. Econ. Manage.*, *34*(3), 240–255, 1997.
Kaoru, Y., Differentiating use and nonuse values for coastal pond water quality improvements, *Environ. Resour. Econ.*, *3*, 487–494, 1993.
Lant, C. L., and R. S. Roberts, Greenbelts in the cornbelt: Riparian wetlands, intrinsic values, and market failure, *Environ. Plann.*, *22*, 1375–1388, 1990.
List, J. A., and C. A. Gallet, What Experimental protocol influence disparities between actual and hypothetical stated values?, *Environ. Resour. Econ.*, *20*(3), 241–254, 2001.
Magat, W. A., J. Huber, K. W. Viscusi, and J. Bell, An Iterative choice approach to valuing clean lakes, rivers, and streams, *J. Risk Uncertainty*, *21*(1), 7–43, 2000.
Mitchell, R. C., and R. T. Carson, An experiment in determining willingness to pay for national water quality improvements, draft report, Resour. for the Future, Washington, D. C., 1981.
Mitchell, R. C., and R. T. Carson, Using surveys to value public goods: The contingent valuation method, report, Resour. for the Future, Washington, D. C., 1989.
Mrozek, J. R., and L. O. Taylor, What determines the value of life? A meta-analysis, *J. Policy Anal. Manage.*, *21*(2), 253–270, 2002.
Olsen, D., J. Richards, and R. D. Scott, Existence and sport values for doubling the size of Columbia River Basin Salmon and Steelhead Runs, *Rivers*, *2*(1), 44–56, 1991.
Poe, G. L., K. J. Boyle, and J. C. Bergstrom, A preliminary meta analysis of contingent values for ground water quality revisited, in *The Economic Value of Water Quality*, edited by J. C. Bergstrom, K. J. Boyle, and G. L. Poe, pp. 137–162, Edward Elgar, Northampton, Mass., 2001.
Poole, C., and S. Greenland, Random-effects meta-analyses are not always conservative, *Am. J. Epidemiol.*, *150*(5), 469–474, 1999.
Ready, R. C., J. C. Buzby, and D. Hu, Differences between continuous and discrete contingent valuation estimates, *Land Econ.*, *72*(3), 397–411, 1996.
Roberts, L. A., and J. A. Leitch, Economic valuation of some wetland outputs of Mud Lake, *Agric. Econ. Rep. 381*, Dep. of Agric. Econ., N. D. State Univ., Fargo, 1997.
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, *J. Agric. Appl. Econ.*, *32*(3), 459–470, 2000a.
Rosenberger, R. S., and J. B. Loomis, Using meta-analysis for benefit transfer: In-sample convergent validity tests of an outdoor recreation database, *Water Resour. Res.*, *36*(4), 1097–1107, 2000b.
Rowe, R. D., W. D. Schulze, B. Hurd, and D. Orr, Economic assessment of damage related to the Eagle Mine Facility, report, Energy and Resour. Consultants, Inc., Boulder, Colo., 1985.
Sanders, L. B., R. G. Walsh, and J. B. Loomis, Toward empirical estimation of the total value of protecting rivers, *Water Resour. Res.*, *26*(7), 1345–1357, 1990.
Santos, J. M. L., *The Economic Valuation of Landscape Change*, Edward Elgar, Northampton, Mass., 1998.
Singer, J. D., Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models, *J. Educ. Behavioral Stat.*, *23*(4), 323–355, 1998.
Smith, V. K., and L. L. Osborne, Do contingent valuation estimates pass a "scope" test? A meta-analysis of visibility improvements at U.S. national parks, *J. Environ. Econ. Manage.*, *31*(3), 287–301, 1996.
StataCorp, Stata Statistical Software: Release 7.0 reference manual, College Station, Tex., 2001.
Sutherland, R. J., and R. G. Walsh, Effect of distance on the preservation value of water quality, *Land Econ.*, *61*(3), 282–291, 1985.
Walsh, R. G., D. A. Greenley, R. A. Young, J. R. McKean, and A. A. Prato, Option values, preservation values and recreational benefits of improved water quality: A case study of the South Platte River Basin, Colorado, *Rep. EPA 600/5 78 001*, U. S. Environ. Prot. Agency, Research Triangle Park, N. C., 1978.
Welle, P. G., Potential economic impacts of acid deposition: A contingent valuation study of Minnesota, Ph.D. dissertation, Univ. of Wis., Madison, 1986.
Whitehead, J. C., and G. C. Blomquist, A link between behavior, information, and existence value, *Leisure Sci.*, *13*, 97–109, 1991.
Whitehead, J. C., and P. A. Groothuis, Economic benefits of improved water quality: A case study of North Carolina's Tar Pamlico River, *Rivers*, *3*, 170–178, 1992.
Whitehead, J. C., G. C. Blomquist, T. J. Hoban, and W. B. Clifford, Assessing the validity and reliability of contingent values: A comparison of on site users, off site users, and non users, *J. Environ. Econ. Manage.*, *29*(2), 238–251, 1995.
Whittington, D., G. Cassidy, D. Amaral, E. McClelland, H. Wang, and C. Poulos, The economic value of improving the environmental quality of Galveston Bay, *Rep. GBNEP-38*, Dep. of Environ. Sci. and Eng., Univ. of N. C., Chapel Hill, 1994.
Wolf, F., *Meta-Analysis: Quantitative Methods for Research Synthesis*, Sage, Beverly Hills, Calif., 1986.
Woodward, R. T., and Y. S. Wui, The economic value of wetland services: A meta analysis, *Ecol. Econ.*, *37*(2), 257–270, 2001.

E. Y. Besedin and R. F. Wardwell, Abt Associates, Inc., 55 Wheeler Street, Cambridge, MA 02138-1168, USA.
R. J. Johnston, Department of Agricultural and Resource Economics, University of Connecticut, Avery Point, Groton, CT 06340, USA. (rjohnston@canr.uconn.edu)