WATER RESOURCES RESEARCH, VOL. 36, NO. 4, PAGES 1097–1107, APRIL 2000

Using meta-analysis for benefit transfer: In-sample convergent validity tests of an outdoor recreation database

Randall S. Rosenberger and John B. Loomis

Department of Agricultural and Resource Economics, Colorado State University, Fort Collins

Abstract. The application of metaregression analysis models for the purpose of benefit transfer is investigated using in-sample convergent validity tests on average value transfers. The database on which the metaregression analysis models are developed is composed of empirical outdoor recreation use value studies conducted from 1967 through 1998. Results of the convergent validity tests suggest that the national model is slightly more robust to changes in application than the Census Region models. The results suggest that the application of meta-analysis for benefit transfers is promising considering limitations imposed by inconsistent data reporting of original studies.

1. Introduction

The formal study of benefit transfer was launched with a special section of *Water Resources Research*, *28*(3), 651–722, 1992. Benefit transfer is the application of values and other information from a "study" site with data to a "policy" site with little or no data. Primary research, while being a "first-best" strategy, is expensive and time-consuming. In other cases, what is at risk at an unstudied site does not justify the high costs of doing primary research. Therefore benefit transfer, as a "second-best" strategy, can be very important in many management contexts. In cases where either strategy is feasible, the decision between the two strategies is then whether the benefit of more information from primary research is worth its additional cost over the less expensive benefit transfer.

Several researchers have investigated the applicability and precision of benefit transfers since 1992. However, these past investigations were primarily concerned with traditional methods of benefit transfer, including point estimate transfer, an averaging of values transfer, expert judgment concerning values transfer, and benefit (demand) function transfer. *Brookshire and Neill* [1992] and *Desvousges et al.* [1998] provide good reviews of the issues in benefit transfer applications. Only more recently has meta-analysis using multivariate regression been investigated for use in benefit transfer [*Desvousges et al.*, 1998; *Kirchhoff*, 1998; L. A. Sturtevant et al., unpublished manuscript, 1998].

Metaregression analysis is the statistical summarizing or synthesizing of past research results using multivariate regression analysis [*Stanley and Jarrell*, 1989]. Meta-analysis has a long history in the fields of psychology, education, and the health sciences. Past meta-analyses have been conducted on recreation activities [*Walsh et al.*, 1989, 1992; *Smith and Kaoru*, 1990; L. A. Sturtevant et al., unpublished manuscript, 1998], groundwater [*Boyle et al.*, 1994], air quality [*Smith and Huang*, 1995; *Smith and Osborne*, 1996; *Desvousges et al.*, 1998], endangered species [*Loomis and White*, 1996], and price elasticities for water [*Espey et al.*, 1997].

Meta-analysis is traditionally concerned with understanding the influence of methodological and study-specific factors on research outcomes and providing statistical summaries and syntheses of past research. A more recent use of meta-analysis is the systematic utilization of the existing value estimates from the literature for the purpose of benefit transfer. Essentially, meta-analysis regression models can be used to "forecast" benefits at unstudied policy sites. Meta-analysis has several conceptual advantages over other benefit transfer methods such as point estimate, average value, and demand function transfers. First, meta-analysis utilizes information from a greater number of studies, thus providing more rigorous measures of central tendency that are sensitive to the underlying distribution of the study values. Second, methodological differences can be controlled for when calculating a value from the meta-analysis function. Third, by setting the independent variables at levels specific to the policy site, the analyst is potentially accounting for differences between the original study sites and the transfer site. Fourth, multiactivity, multisite meta-analyses can provide estimates for regions in which no studies were conducted for an activity. That is, meta-analysis can project estimates for new or unstudied activities.

There are also empirical limitations to using meta-analysis for benefit transfer [*Desvousges et al.*, 1998]. First, there should be enough original studies conducted so that statistical inferences can be made. Second, a meta-analysis can only be as good as the quality of past research efforts. This quality includes the scientific soundness of the original research and the reporting of results and summary statistics on the original data sample that is rich in detail. Third, the studies should be similar enough in content and context that they can be combined and statistically analyzed.

One of the most important issues for benefit transfer is how well it performs. Does it provide valid measures of benefits at unstudied sites? One approach to answering this question is to conduct convergent validity tests. There have been essentially two types of convergent validity tests [*Desvousges et al.*, 1998]. The first type compares a transfer value to the "true" value for a site, where the true value comes from an original study at the site. However, it should be noted that this true value is itself an estimate of the unknown value. The second type of convergent validity test is to compare two transfer estimates to determine if the method applied is invariant to judgments by the analyst. The results from both of these types of convergent validity tests can then be used to anchor our confidence in actually applying benefit transfers to unstudied sites.

Paper number 2000WR900006.
0043-1397/00/2000WR900006$09.00

Several studies have performed convergent validity tests on benefit transfer trials using different transfer methods [*Loomis*, 1992; *Smith and Huang*, 1995; *Loomis et al.*, 1995; *Downing and Ozuna*, 1996; *Kirchhoff et al.*, 1997; *Desvousges et al.*, 1998; *Kirchhoff*, 1998]. While the evidence provides some confidence in pursuing benefit transfers, with several cases producing values very similar to the true values (as low as a few percentage points), in other cases the disparity between the true value and the transfer value is quite large (in excess of 800%).

Kirchhoff [1998] illustrates the basic use of meta-analysis for benefit transfer. She uses a meta-analysis model estimated by *Walsh et al.* [1989, 1992] and benefit function models estimated by *Kirchhoff et al.* [1997] to predict values for a policy site. She then conducts convergent validity tests comparing multisite benefit function transfers, single-site benefit function transfers, and meta-analysis-based transfers to the true or primary value of the site. She found that multisite benefit functions outperformed meta-analysis but that meta-analysis outperformed single-site benefit function transfers. *Kirchhoff* [1998] concludes that although meta-analysis was outperformed by multisite benefit function transfers, meta-analysis held up well considering the bias of her methodology toward benefit function transfers. Also, note that the *Walsh et al.* [1989, 1992] meta-analysis was not designed for benefit transfer purposes. L. A. Sturtevant et al. (unpublished manuscript, 1998) support *Kirchhoff's* [1998] conclusion by showing that in general, estimates from a meta-analysis are more precise than point estimate transfers.

The purpose of this paper is to contribute to the refinement and testing of meta-analyses as a benefit transfer tool. We first update the national meta-analysis of *Walsh et al.* [1989, 1992] with the last decade of research on outdoor recreation use values. We then optimize national and geographic-region-specific meta-analysis models for use in benefit transfers. We then evaluate the relative accuracy of these benefit transfer models using in-sample convergent validity tests. These in-sample tests provide an indication of the robustness of the meta-analysis benefit transfer models to different applications. The remainder of this paper will present the data, the meta-analysis regressions and benefit transfer models, results of the in-sample convergent validity tests, with a focus on applying one of the models to benefit transfer for water-related activities.

2. Meta-analysis Design and Results

2.1. Data

The data used for this analysis are based on two extensive literature reviews of outdoor recreation use value studies. The first literature review was conducted by *Walsh et al.* [1988] spanning 1967–1988. This review was extensively coded by *MacNair* [1993], enabling its integration with the second review. The second literature review coded outdoor recreation use value studies spanning the period from 1989 to 1998 (R. S. Rosenberger et al., unpublished working paper, 2000). One of the main objectives of the second literature review was to target recreation activities that were either unrepresented or underrepresented in the first literature review. Thus heavily studied activities such as fishing and big game hunting were not emphasized in the second review; these activities being well represented by the previous literature review.

All study values were adjusted to per person activity day units and updated from their original study year (not publication year) values to 1996 dollars using the Implicit Price Deflator. Originally there were slightly more than 170 individual studies that produced slightly more than 750 individual values. Some of these studies were removed from the database because they did not report enough information to adjust their reported values to a per day basis. Other studies were not included in the meta-analysis because of the lack of reporting of key information that would enable the full coding of a study. Therefore the meta-analysis database consists of 682 estimates from 131 separate studies.

Table 1 presents a summary of the use value estimates per activity by U.S. Bureau of Census Region (CR), estimates that are national in scope, and estimates for Canada for those studies included in the meta-analysis (a complete bibliography of individual studies and reported values is available from the authors upon request). We combined CR5 (Alaska) with CR4 (the Pacific states) based on Chow test results to form CR45 (the Pacific Coast), thus preserving degrees of freedom in the meta-analysis benefit transfer modeling and mirroring the U.S. Forest Service's Resource Planning Act (RPA) assessment regions. *Walsh et al.* [1989, 1992] reported their data segregated by U.S. Forest Service Regions (the sponsoring agency). This segregation results in two problems: (1) very small sample sizes per activity/region cell and (2) numerous activity/region cells with no average value (because of the lack of any original studies). To address both of these problems, the U.S. Forest Service Regions were aggregated into Census Regions. The average values reported in Table 1 will serve as the benchmark or true values we will attempt to forecast using the meta-analysis benefit transfer models. It is a common practice in federal public land agencies to use agency-approved average values in management and policy assessments. The U.S. Forest Service has done this since 1980 using their Resources Planning Act (RPA) values [*U.S. Forest Service*, 1989]. The RPA values are specific to groups of similar activities and regions of the country. The U.S. Bureau of Reclamation and U.S. Army Corps of Engineers have relied upon the U.S. Water Resources Council's Unit Day Values [*U.S. Water Resources Council*, 1973, 1979, 1983] for decades. Therefore we will assess the convergent validity of meta-analysis-forecasted average values to these true average values for different recreation activities for each Census Region.

A master coding sheet was developed that contains 126 fields. The main coding categories include reference, benefit measure(s), methodology used, recreation activity investigated, recreation site characteristics, and user or sample population characteristics. Table 2 lists and defines the variables tested in the meta-analysis of different factors on use value estimates. The majority of these variables are qualitative dummy variables coded as 0 or 1, where 0 means the study does not have a characteristic and 1 means that it does. For example, if an open-ended technique was used to elicit value information from respondents, then both METHOD (for stated preference model type) and OE would be coded as 1, while other mutually exclusive variables would be coded 0. Table 2 groups the variables according to whether they are methodological (including revealed preference (RP) and stated preference (SP) types and subtypes, survey mode, and functional form specification) or site- (including geographic location based on Forest Service (FS) Regions and site characteristics) or activity-specific variables.

The user population characteristics were rarely reported with the results of a study. Other means for obtaining data on user population characteristics, such as contacting the researchers of a study, were not feasible given the financial and

Table 1. Benchmark Raw Average Recreation Values per Person per Activity Day

	CR1: North		CR2: South		CR3: Rocky Mountains		CR45: Pacific Coast		United States		Canada	
Activity	N	Value	N	Value	N	Value	N	Value	N	Value	N	Value
Water-related												
Swimming	5	37.21					1	14.95	1	20.67		
Float boating			2	8.40	4	72.42	4	21.69	1	21.61		
Motorboating	2	52.44			2	68.76	1	15.13	1	38.70		
Waterfowl hunting	23	32.09	11	17.70	19	36.74	6	30.44				
Fishing	42	31.63	13	27.74	39	42.49	16	37.11	4	37.26	4	55.15
Non-water-related												
Camping	3	30.07	6	20.35	10	22.42	3	113.88	1	28.61		
Picnicking	1	55.22	1	37.24	1	32.30	2	73.95	1	15.69		
Sightseeing	1	129.04	4	42.54	7	36.31	1	50.64	1	18.83		
Off-road driving							1	33.64	1	19.94		
Hiking/backpacking	2	45.01	2	109.96	3	37.42	6	20.39	1	20.87		
Biking	1	34.11	1	56.27	2	58.89			1	17.61		
Downhill skiing					2	23.23	1	20.90	1	19.61		
Cross-country skiing	2	28.83			1	11.71			1	13.20		
Snowmobiling					1	36.23						
Big game hunting	55	45.22	26	35.99	68	45.05	16	46.47	2	104.90	3	35.60
Small game hunting	3	36.73			13	25.75			2	103.02		
Wildlife viewing	56	27.03	38	29.50	38	32.80	22	41.72				
Horseback riding									2	16.08		
Rock climbing	2	85.74			3	42.04						
General recreation	8	15.21	6	14.65	17	31.82	15	15.91			3	3.33
Other			2	26.21	8	52.36						
Total number of cases	206		112		238		95		21		10	

N, number of estimates (total is 682). Recreation values are given in U.S. dollars.

time constraints of the project. We did attempt to proxy user population characteristics by using 1990 U.S. Census average values for income, gender, education, age, and race for the state in which the study was conducted, but found in preliminary analysis that these proxies were broadly insensitive to differences in the use values provided. Using U.S. Census average values from the period closest to when the original study data were collected may be a viable alternative in future investigations.

2.2. Meta-analysis Models

The meta-analysis models are of the basic form:

$$y_{im} = \alpha + \beta' x_{im} + \varepsilon_m, \tag{1}$$

where i indexes each observation, m indexes whether the data are for a Census Region model or the national model, y is the dependent variable (in this case, consumer surplus per person day adjusted to 1996 dollars), α and β are parameters to be estimated and are the intercept and slope coefficients for the models, respectively, x is a vector of explanatory variables including methodology, site, and activity characteristics, and ε is the classical error term with mean zero and variance σ_ε^2. The range of the dependent variable is from a low of \$1 to a high of \$218, with a mean of \$36 and median of \$28. Figure 1 displays the distribution of the dependent variable.

The 682 observations included in the meta-analyses were provided from 131 separate studies. The number of estimates per study ranged from 1 to 134. As identified in previous meta-analyses, the panel nature of the data can lead to econometric problems. If there is correlation among these multiple observations for each study, then a classical ordinary least squares (OLS) regression will be inefficient and inconsistent in estimated parameters. We tested for panel effects using various forms of stratifying the data (including the most obvious stratification by study) (R. S. Rosenberger and J. B. Loomis, unpublished working paper, 2000). However, panel effects were not discernible with these tests. Therefore we use classical ordinary least squares with the Newey-West version of White's consistent covariance estimator to estimate the models [*Smith and Kaoru*, 1990]. There is no precedent for choice of functional form when conducting meta-analyses. The functional form of the meta-analysis models is linear in the dependent variable and the quantitatively defined variables, with the majority of the variables being qualitative dummy variables as noted in section 2.1. We tested and rejected logarithmic transformations of the quantitatively defined variables, finding the linear specification to be most efficient.

2.3. Results

Six meta-analysis models were estimated, a fully specified model for traditional investigations of factor effects on study outcomes and five models optimized for benefit transfer applications. The benefit transfer meta-analysis models include a national (NAT) model and one for each of the Census Region (CR)–defined geographic regions (the CR models are labeled as CR1 (north), CR2 (south), CR3 (Rocky Mountains), and CR45 (Pacific coast).

2.3.1. Traditional meta-analysis model. Table 3 presents the traditional meta-analysis model that investigates the systematic effects of the different variables identified in Table 2. The model has an adjusted R^2 of 0.26, slightly below that of *Walsh et al.* [1989, 1992] for their combined SP and RP model. Although the model is overspecified and there are problems of correlation among some of the independent variables, several variables were significant in the model at the 90% level or better. Elsewhere a more detailed explanation of the traditional model is provided (R. S. Rosenberger et al., unpublished working paper, 2000).

Table 2. Description of Variables in the Meta-analysis Models

Variable	Description
Dependent variable	
CS	consumer surplus (CS) per person per activity day (1996 dollars)
Method variables	
METHOD	qualitative variable: 1 if stated preference (SP) valuation approach used; 0 if revealed preference (RP) approach used
DCCVM	qualitative variable: 1 if SP and dichotomous choice elicitation technique was used; 0 if otherwise
OE	qualitative variable: 1 if SP and open-ended elicitation technique was used; 0 if otherwise
ITBID	qualitative variable: 1 if SP and iterative bidding elicitation technique was used; 0 if otherwise
PAYCARD	qualitative variable: 1 if SP and payment card elicitation technique was used; 0 if otherwise
CONJOINT	qualitative variable: 1 if SP and conjoint analysis technique was used; 0 if otherwise
RPSP	qualitative variable: 1 if SP and RP used in combination; 0 if otherwise
ZONAL	qualitative variable; 1 if RP and a zonal travel cost model was used; 0 if otherwise
INDIVID	qualitative variable: 1 if RP and an individual travel cost model was used; 0 if otherwise
RUM	qualitative variable: 1 if RP and a random utility model was used; 0 if otherwise
HEDONIC	qualitative variable: 1 if RP and a hedonic travel cost model was used; 0 if otherwise (omitted category for METHOD) [0.024, 0.15]*
TTIME	qualitative variable: 1 if RP model included travel time; 0 if otherwise
SUBS	qualitative variable: 1 if RP model included substitute sites; 0 if otherwise
ONSITE	qualitative variable: 1 if sample frame was on site; 0 if otherwise
MAIL	qualitative variable: 1 if primary data collection used mail survey type; 0 if otherwise
PHONE	qualitative variable: 1 if primary data collection used phone survey type; 0 if otherwise
INPERSON	qualitative variable: 1 if primary data collection used in-person survey type; 0 if otherwise
SECOND	qualitative variable: 1 if secondary data were used (omitted category for data collection) [0.063, 0.24]*
LINLIN	qualitative variable: 1 if functional form was linear on both dependent (d.v.) and independent variables (i.v.); 0 if otherwise
LOGLIN	qualitative variable: 1 if functional form was log d.v. and linear i.v.; 0 if otherwise
LOGLOG	qualitative variable: 1 if functional form was log on both d.v. and i.v.; 0 if otherwise
LINLOG	qualitative variable: 1 if functional form was linear on d.v. and log on i.v.; 0 if otherwise (omitted category for functional form) [0.003, 0.05]*
VALUNIT	qualitative variable: 1 if CS was originally estimated as per day; 0 if otherwise (e.g., trip, season, or year)
TREND	qualitative variable: year when data were collected, coded as 1967 = 1, 1968 = 2,..., 1996 = 30
Site variables	
RECQUAL	qualitative variable: site quality variable coded as 1 if the author stated site was of high quality or the site was either a National Park, National Recreation Area, or Wilderness Area; 0 if otherwise
FSADMIN	qualitative variable: 1 if the study site(s) were National Forests (i.e., administered by the U.S. Forest Service (FS)); 0 if otherwise
R1	qualitative variable: 1 if study sites were in FS Region 1 (Montana, northern Idaho); 0 if otherwise
R2	qualitative variable: 1 if study sites were in FS Region 2 (Wyoming, Colorado); 0 if otherwise
R3	qualitative variable: 1 if study sites were in FS Region 3 (Arizona, New Mexico); 0 if otherwise
R4	qualitative variable: 1 if study sites were in FS Region 4 (Nevada, Utah, southern Idaho); 0 if otherwise
R5	qualitative variable: 1 if study sites were in FS Region 5 (California); 0 if otherwise
R6	qualitative variable: 1 if study sites were in FS Region 6 (Oregon, Washington); 0 if otherwise
R8	qualitative variable: 1 if study sites were in FS Region 8 (southern United States east of Rocky Mountains); 0 if otherwise
R9	qualitative variable: 1 if study sites were in FS Region 9 (northern United States east of Rocky Mountains); 0 if otherwise
R10	qualitative variable: 1 if study sites were in FS Region 10 (Alaska); 0 if otherwise
NATL	qualitative variable: 1 if study sites were the entire United States; 0 if otherwise
CANADA	qualitative variable: 1 if study sites were in Canada; 0 if otherwise (omitted category for geographic location of study site) [0.015, 0.12]*
LAKE	qualitative variable: 1 if the recreation site was a lake; 0 if otherwise
RIVER	qualitative variable: 1 if the recreation site was a river; 0 if otherwise
FOREST	qualitative variable: 1 if the recreation site was a forest; 0 if otherwise
OCEAN	qualitative variable: 1 if the recreation site was an estuary or bay of an ocean; 0 if otherwise (omitted category for site type) [0.169, 0.37]*
PUBLIC	qualitative variable: 1 if ownership of the recreation site was public; 0 if otherwise.
DEVELOP	qualitative variable: 1 if the recreation site had developed facilities, such as picnic tables, campgrounds, restrooms, boat ramps, ski lifts, etc.; 0 if otherwise.
NUMACT	quantitative variable: the number of different recreation activities the site offers.
Recreation activity variables	
CAMP... OTHERREC	qualitative variables: 1 if the relevant recreation activity was studied; 0 if otherwise. CAMP is camping, PICNIC is picnicking, SWIM is swimming, SISEE is sightseeing, OFFRD is off-road driving, NOMTRBT is float boating, MTRBOAT is motorboating, HIKE is hiking/backpacking, BIKE is biking, DHSKI is downhill skiing, XSKI is cross-country skiing, SNOWMOB is snowmobiling, BGHUNT is big game hunting, SMHUNT is small game hunting, WATFOWL is waterfowl hunting, FISH is fishing, WLVIEW is wildlife viewing, HORSE is horseback riding, ROCKCL is rock climbing, GENREC is general recreation (defined as a composite of recreation activity opportunities at a site), and OTHERREC is other recreation (for sites with recreation opportunities undefined or obscure: omitted category for recreation activity) [0.015, 0.12]*

CS, consumer surplus; SP, stated preference; RP, revealed preference; d.v. and i.v., dependent and independent variables, respectively; FS, Forest Service.

*Mean and standard deviation for omitted categories are reported in square brackets; $N = 682$.

METHOD is negative and significant in the model, meaning that SP models yield lower estimates than RP models in general, which is consistent with the *Walsh et al.* [1989, 1992] meta-analysis and the bulk of travel cost-contingent valuation model comparison studies [*Carson et al.*, 1996]. The use of a zonal travel cost (ZONAL) technique in RP models is significant and yields lower estimates than other valuation techniques. The inclusion of substitute sites (SUBS) in RP models was negative and significant, indicating that the better job a RP model does of reflecting substitutes, the lower the benefit estimate will be [*Rosenthal*, 1987]. The use of a phone survey mode (PHONE) is negative and significant, implying that phone surveys yield relatively lower estimates than other survey modes. The variable for original value units reported (VALUNIT) is significant and negative in the model, implying that estimates originally reported in per day units are lower than per day estimates converted from values originally reported as per trip, per season, or per year.

The TREND variable is significant and positive, meaning there has been an increase in benefit measures in real terms over time. U.S. Forest Service–administered lands (FSADMIN) yield lower consumer surplus values in our model, which may be because of Forest Service–administered sites being identified in juxtaposition to sites of designated higher quality (e.g., National Parks, State Parks, and National Wildlife Refuges). Lake recreation (LAKE) yields lower consumer surplus values, and river recreation (RIVER) yields higher consumer surplus values, than ocean or bay recreation. Recreation on public lands (PUBLIC), in general, provides higher consumer surplus values than private areas. A plausible reason for this result is that private areas charge substantially more for access and onsite facilities and services than public areas. This means that people who use private areas have more of their consumer surplus value extracted through fees than people who use public areas, where lower and/or fewer fees are charged. Big game hunting (BGHUNT), waterfowl hunting (WATFOWL), fishing (FISH), and rock climbing (ROCKCL) were all significant and positive in the model, yielding discernibly higher consumer surplus values than other recreation activities in general.

2.3.2. Benefit transfer meta-analysis models. The meta-analysis benefit transfer models were optimized by retaining only those variables that were significant at an 80% level of confidence or better based on t-statistics. This optimization is necessary in order to account for region-specific differences and to reduce overspecification of the models in retaining variables whose coefficients are not significantly different than zero. Beyond some point, having too may independent variables makes use of meta-analysis functions for benefit transfer more cumbersome. A stepwise procedure was used to optimize each benefit transfer model. The procedure began with the full specification of the traditional model shown in Table 3 and sequentially eliminated the least significant variable until all remaining variables are significant at the 80% confidence level or better (p value equals 0.20).

Table 4 presents the optimized meta-analysis models for benefit transfer applications. The adjusted R^2 for the optimized benefit transfer NAT model compared to the traditional meta-analysis model remains at 0.26. The CR models all had higher adjusted R^2 than the NAT model, including 0.28 for CR1, 0.66 for CR2, 0.36 for CR3, and 0.33 for CR45. The signs and significance of the variables in the NAT model are consistent with the traditional model, with additional variables being significant in the NAT model as compared to the fully

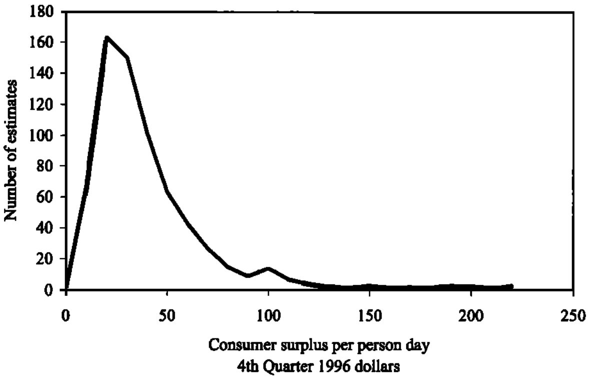

Figure 1. Distribution of dependent variable, $N = 682$.

specified traditional model. Direct comparisons cannot be made between the CR models and the NAT and traditional models because of differences in sample composition. The CR models are developed using data specific to the region.

None of the variables are significant across all of the benefit transfer models with the exception of big game hunting (BGHUNT). The majority of the methodological variables have consistent signs when significant in two or more models. About half of the site variables have consistent signs when significant in two or more models. However, the majority of the recreation activity variables have inconsistent signs when significant in two or more models. These results are reasonable when we consider that as variables are dropped during the variable selection procedure, the implied composite constant term changes. Therefore each coefficient represents an incremental change of that variable relative to different composite bases. For example, in the national model the activity portion of the composite base begins with OTHERREC and ends with OTHERREC/CAMP/PICNIC/SISEE/DHSKI/XSKI/WLVIEW/GENREC. Therefore interpretations of the remaining activity variable coefficients are now incremental changes from the latter and not the former as in the traditional model. Each benefit transfer model has a different composite base.

These benefit transfer functions can be used to forecast use values sensitive to methodology, recreation activity, and site characteristics. Forecasting is accomplished by multiplying each coefficient on an independent or explanatory variable by some value and then aggregating the effects. In the case of qualitative variables they can be either "turned on" (a value of 1 providing the full effect of the variable), "turned off" (a value of 0 where the variable has no effect), or some value in between 0 and 1. For example, if you wanted to forecast the value of fishing, the coefficient on FISH would be turned on (multiplied by 1) and all other activity variable coefficients would be turned off (multiplied by 0). In the case of quantitative variables some other criteria would be used to determine the magnitude of the variable effect to be included, such as the current year for TREND or the average number of activities offered at the policy site for NUMACT.

3. Convergent Validity Tests and Results

Recently, the U.S. Forest Service has become interested in the possibility of conducting benefit transfers using meta-analysis derived models. In addition, benefit transfer meta-analysis can be used to at least fill in the missing values in an RPA value by recreation activity by region table (e.g., the blank

Table 3. Meta-analysis Regression of Outdoor Recreation Use Value Studies: 1967–1998

Variable	Coefficient	Uncorrected Standard Error of Coefficient	White's Consistent Covariance Standard Error of Coefficient*	Mean of Variable	Standard Deviation of Variable
CS	dependent variable	...	...	36.10	32.39
METHOD	−36.953†	17.39	15.63	0.64	0.48
DCCVM	3.9405	12.82	11.03	0.18	0.38
OE	−7.9101	12.52	9.08	0.36	0.47
ITBID	−2.7065	12.42	8.54	0.01	0.30
PAYCARD	−34.246†	23.45	17.23	0.01	0.05
SPRP	−31.564†	19.89	14.37	0.01	0.08
ZONAL	−34.999†	18.37	18.85	0.20	0.40
INDIVID	−21.354	18.48	20.48	0.14	0.34
RUM	−28.637	18.60	18.27	0.03	0.16
TTIME	3.1637	6.67	11.22	0.31	0.46
SUBS	−21.304†	5.08	9.00	0.26	0.44
ONSITE	5.4999	8.93	13.12	0.29	0.46
MAIL	−7.0068	5.18	4.29	0.25	0.43
PHONE	−20.922†	5.41	5.82	0.51	0.50
INPERSON	−6.0011	6.72	8.52	0.35	0.48
LINLIN	8.7587	8.06	9.31	0.10	0.31
LOGLIN	7.0837	7.84	8.49	0.16	0.36
LOGLOG	3.4291	8.76	11.60	0.06	0.24
VALUNIT	−10.061†	3.98	4.92	0.39	0.49
TREND	1.2460†	0.37	0.57	19.04	5.33
RECQUAL	9.3201	6.11	13.02	0.11	0.31
FSADMIN	−15.209†	4.46	5.48	0.14	0.34
R1	10.873	11.59	14.82	0.05	0.22
R2	−3.3803	10.90	12.95	0.12	0.32
R3	−0.6597	11.62	14.02	0.06	0.24
R4	2.1742	11.04	13.77	0.11	0.32
R5	−6.0734	11.63	13.64	0.05	0.22
R6	−13.562	11.52	13.04	0.06	0.24
R8	−6.2976	10.87	12.68	0.16	0.37
R9	−4.3382	10.74	13.20	0.30	0.46
R10	−8.9159	12.72	13.10	0.03	0.16
NATL	14.771	15.20	33.54	0.03	0.17
LAKE	−12.973	7.09	9.40	0.05	0.22
RIVER	21.117†	6.87	8.87	0.04	0.20
FOREST	−6.2761	4.50	6.87	0.30	0.46
PUBLIC	19.951†	8.50	10.02	0.96	0.20
DEVELOP	−5.5700	6.33	9.08	0.19	0.39
NUMACT	−0.1040	0.39	0.65	4.64	9.08
CAMP	17.611	12.20	14.07	0.03	0.18
PICNIC	12.327	16.12	13.37	0.01	0.09
SWIM	6.0168	15.15	16.71	0.01	0.10
SISEE	11.043	14.22	19.88	0.08	0.27
OFFRD	4.4532	22.95	12.82	0.01	0.05
NOMTRBT	12.606	15.68	14.33	0.01	0.09
MTRBOAT	17.092	13.38	16.32	0.02	0.13
HIKE	7.6269	13.43	18.06	0.02	0.14
BIKE	−5.2075	17.59	14.28	0.01	0.08
DHSKI	22.778	18.45	17.34	0.01	0.08
XSKI	8.8033	17.97	11.96	0.01	0.08
SNOWMOB	1.7078	31.81	14.82	0.01	0.04
BGHUNT	27.866†	11.19	11.46	0.25	0.43
SMHUNT	19.674	12.84	12.81	0.03	0.16
WATFOWL	23.145†	11.49	11.49	0.09	0.28
FISH	20.982†	10.73	11.24	0.17	0.38
WLVIEW	12.605	11.32	12.65	0.16	0.37
HORSE	−1.2599	31.18	15.69	0.01	0.04
ROCKCL	60.908†	19.85	14.63	0.01	0.08
GENREC	14.119	12.81	18.22	0.07	0.26
CONSTANT	37.239	25.86	31.15	...	...
Adjusted R^2	0.26				
F-statistic (degrees of freedom)	5.109† (58, 623)				

$N = 682$.

*Standard errors of parameter estimates corrected for heteroskedasticity and serial correlation using the Newey-West version of White's consistent covariance estimator and 11 periods.

†Variable coefficient is statistically significant at the 0.10 level or better based on t-statistics from corrected standard errors of the estimated parameters.

Table 4. Optimized Benefit Transfer Meta-analysis Models

Variable	National Coefficient	CR1 Coefficient	CR2 Coefficient	CR3 Coefficient	CR45 Coefficient
CONSTANT	19.159*	†	26.541†	66.234*	10.044†
	(10.20)		(24.89)	(15.19)	(10.65)
METHOD	−17.598*			−34.381*	
	(6.81)			(12.17)	
DCCVM				27.066*	
				(13.63)	
OE	−7.4684*				
	(3.82)				
RPSR	−38.170*	na	na		na
	(10.73)				
PAYCARD	−28.333	na	na		na
	(19.72)				
INDIVID		38.927*	52.653*		
		(14.40)	(14.21)		
RUM		63.278*		−28.392*	30.445*
		(29.50)		(16.12)	(14.65)
SUBS	−20.277*			−33.481*	−18.851
	(5.08)			(17.75)	(12.67)
MAIL			19.309		
			(12.01)		
PHONE	−18.626*				
	(4.24)				
LOGLIN					22.642*
					(11.84)
LOGLOG		24.305*			
		(12.11)			
VALUNIT	−5.8206	−18.462*		−10.023*	
	(4.12)	(9.00)		(5.85)	
TREND	1.6132*				
	(0.38)				
FSADMIN	−20.056*		−18.471	−14.595*	
	(4.23)		(11.36)	(8.74)	
R2	−6.5813	na	na	−8.7916	na
	(4.13)			(5.57)	
R5	−10.448	na	na	na	
	(6.96)				
R6	−14.218*	na	na	na	
	(4.55)				
R8	−8.7566*	na	pc	na	na
	(3.74)				
R9	−7.1241*	pc	na	na	na
	(3.83)				
R10	−14.980*	na	na	na	
	(8.50)				
LAKE	−16.803*	−30.097*			
	(6.61)	(15.26)			
RIVER	17.747*		−73.951*	40.462	
	(8.05)		(26.40)	(26.84)	
FOREST			17.792*	−18.358*	
			(3.87)	(10.29)	
PUBLIC	21.655*	29.652*	48.940*		pc
	(5.66)	(6.18)	(27.78)		
DEVELOP			−65.216*		
			(24.30)		
NUMACT				2.2671*	
				(0.72)	
CAMP					107.59*
					(33.30)
PICNIC			−25.683*	−45.120*	60.118*
			(12.30)	(15.82)	(33.37)
SISEE		78.925*	−50.590*		36.809*
		(16.56)	(27.36)		(10.05)
OFFRD	−7.8984	na	na	na	19.803*
	(5.08)				(10.05)
BIKE	−13.569*	−58.772*	−25.962		na
	(7.63)	(19.11)	(15.92)		
DHSKI		na	na	40.033*	
				(17.10)	
XSKI		14.005*	na		na
		(8.20)			
SNOWMOB	−20.299*	na	na	20.026	na
	(9.74)			(12.38)	

Table 4. (continued)

Variable	National Coefficient	CR1 Coefficient	CR2 Coefficient	CR3 Coefficient	CR45 Coefficient
BGHUNT	12.478*	21.70	−48.391*	19.070	34.536
	(3.42)	(13.94)	(23.82)	(11.68)	(22.25)
WATFOWL	10.161*	14.479*	−57.781*		17.827*
	(4.25)	(8.24)	(23.49)		(8.57)
FISH	9.0575*		−61.378*		19.419*
	(4.12)		(25.07)		(7.81)
WLVIEW			−49.923*		30.304*
			(23.42)		(15.94)
ROCKCL	39.738*		na	28.222*	na
	(12.59)			(13.57)	
HORSE	−11.841*	na	na	na	na
	(5.11)				
GENREC					24.721*
					(13.74)
Adjusted R^2	0.26	0.28	0.66	0.36	0.33
F-statistic (degrees of freedom)	9.98* (26, 655)	8.89* (10, 195)	16.14* (14, 97)	10.08* (15, 222)	4.84* (12, 82)
Sample size	682	206	112	238	95

Abbreviations are as follows: na, not applicable or no observations for particular model; pc, perfectly correlated with regional demarcation. Standard errors in parentheses are corrected for heteroskedasticity and serial correlation using the Newey-West version of White's covariance consistent estimator and 11 periods. Dependent variable is consumer surplus per person per activity day in 1996 dollars.

*Here $p < 0.10$ (all variables are $p \leq 0.20$, with the exception of the constant terms as noted below.

†The constant or intercept term for CR1 was dropped ($p > 0.90$), while we retained the constant for CR2 ($p = 0.29$) and CR45 ($p = 0.34$).

cells in Table 1). An indication of the relative accuracy and resultant level of confidence in using meta-analysis for benefit transfer can be determined by performing in-sample convergent validity tests. These tests assess the precision with which the meta-analysis can be used to predict the activity values per region where they exist.

3.1. In-Sample Convergent Validity Tests

The in-sample convergent validity tests are based on adapting the metamodels to regional and activity-specific conditions by substituting the region and activity values for the explanatory variables. This approach provides a predicted average value based on the regression model. The predicted average values for the jth activity in the rth region using the mth model are

$$\hat{\bar{y}}^r_{jm} = \alpha + \beta'\bar{x}^r_m. \quad (2)$$

The $\bar{x}^r_m$ values are mean values for the explanatory variables based on geographic region (national mean values and Census Region mean values). We tested other possible mean values, such as national values for methodological variables and regional values for site variables, but found that these other treatments performed relatively worse than the treatments presented here. National mean values are reported in Table 3. Census Region mean values are available from the authors upon request.

The statistics we use to assess the convergent validity of applying the benefit transfer meta-analysis models are percent difference calculations varying on levels of aggregation. The least aggregated statistic is an activity per region percent difference (Δy) and is defined as

$$\Delta y^r_{jm} = \frac{\hat{\bar{y}}^r_{jm} - y^r_j}{y^r_j}, \quad (3)$$

where the left-hand side of (3) is the percent difference for the jth activity in the rth region using the mth model. For each treatment of the models this is the percent difference of the predicted value for an activity in a region from the true value (Table 1) for that activity in that region. These single percent difference statistics are then aggregated to an activity percent difference statistic.

The second statistic aggregates at the activity level, defining an activity percent difference statistic (Δa) for each model as

$$\Delta a^r_j = \overline{\Delta y^r_{jm}} = \frac{\sum |\Delta y^r_{jm}|}{k}. \quad (4)$$

That is, Δa is the average of the absolute value of the Δy for activity j across the different regions r, and k is the number of Δy for activity j. The most aggregated statistic, a grand percent difference statistic (ΔP), is on the model level and is defined as the average of the Δa:

$$\Delta P = \overline{\Delta a^r_j} = \frac{\sum \Delta a^r_j}{20}, \quad (5)$$

where 20 is the number of Δa, or activity values, in the aggregate (there is no true value for horseback riding in the regions). Here ΔP is the statistic that represents the relative accuracy of the models' prediction of the benchmark average values, providing an indication of the degree of certainty for forecast values using the meta-analysis functions.

We report the results from two adaptations of the models by way of holding the x variables at some mean value with the exception of activity variables that are either turned on (1) or off (0) when relevant. The mean values used are national mean values (i.e., the mean value of the x for the entire data set) and Census Region mean values (i.e., the mean value of the x for each regional subset of the data). Therefore the national model will predict average activity values per region by (1) using national mean values ($\mathrm{NAT_{NAT}}$) and (2) using respective Census Region mean values ($\mathrm{NAT_{CR}}$). The Census Region models will predict average activity values per respective region by (3) using national mean values (CR_{NAT}) and (4) using respective Census Region mean values (CR_{CR}).

3.2. Results

Table 5 reports aggregate statistics for each treatment including the range for the Δy_j and Δa_j, ΔP, and a 95% confidence interval for each ΔP. Table 6 reports the Δy_j and Δa_j for each adaptation of the NAT model. The NAT_{NAT} and CR_{NAT} treatments of the respective models are directly comparable since they both use national mean values of the explanatory variables to predict average Census Region values for the different activities. The national model has a grand percent difference (ΔP) of 54% whereas the CR models combine for a ΔP of 71%. For a specific activity in a region the national model's predicted values ranged from −73 to 319% of the benchmark values. The CR models, combined, had predicted values that ranged from −189 to 722% of the benchmark values. One plausible explanation for the smaller range of the NAT model compared to the CR models is that the mean values used to adapt the regression models are based on national averages for each activity, which is the level of development for the NAT model. More variability is introduced into the CR models because the national mean values are not sensitive to CR model differences.

The NAT_{CR} and CR_{CR} treatments of the respective models are directly comparable since they both use Census Region mean values of the explanatory variables to predict average Census Region values for the different activities. The NAT model has a grand percent difference (ΔP) of 47% whereas the CR models combine for a ΔP of 67%. For a specific activity in a region the NAT model's predicted values ranged from −79 to 201% of the benchmark values. The CR models, combined, had predicted values that ranged from −192 to 774% of the benchmark values. We expected the CR models to perform better under this treatment than the NAT model given that this treatment used Census Region–specific mean values for the explanatory variables. However, the results of the tests suggest that the NAT model is more robust to perturbations of the explanatory variables than are the CR models.

The two treatments of the NAT model are fairly close in performance, whether the explanatory variables are held constant at the national or Census Region levels. The CR models also perform similarly between the two treatments. Although the ΔP for each treatment are not significantly different, this aggregate statistic hides the variability of each treatment at the microlevel. The NAT model outperforms the CR models in terms of reducing the chances of large errors in benefit transfer as evidenced by the ranges on the Δy_j. Therefore one can be more confident using the NAT model for benefit transfers.

Table 5. Summary of Convergent Validity Test Results

Treatment	Δy Range, %	Δa Range, %	ΔP, %	ΔP, % (95% CI)
NAT_{NAT}	−73–319	11–143	54	39–68
NAT_{CR}	−79–201	5–102	47	35–59
CR_{NAT}	−189–772	13–291	71	41–102
CR_{CR}	−192–774	6–293	67	37–97

CI, confidence interval. Here Δy is the percent difference of the predicted average value from the raw average value for an activity in a region, Δa is the average of the absolute value of the Δy for an activity, and ΔP is the average of the Δa across all activities.

Table 6. Activity per Region Percent Difference (Δy_j) Results for the National Model

Activity/Model	CR1	CR2	CR3	CR45	Δa
CAMP					
NAT_{NAT}	14	73	60	−69	54
NAT_{CR}	−6	24	48	−71	37
PICNIC					
NAT_{NAT}	−38	−6	11	−53	27
NAT_{CR}	−48	−32	3	−56	35
SWIM					
NAT_{NAT}	−7	na	na	133	70
NAT_{CR}	−24	na	na	119	71
SISEE					
NAT_{NAT}	−73	−17	−1	−31	31
NAT_{CR}	−78	−40	−8	−35	40
OFFRD					
NAT_{NAT}	na	na	na	−20	20
NAT_{CR}	na	na	na	−26	26
NOMTRBT					
NAT_{NAT}	na	319	−50	60	143
NAT_{CR}	na	201	−54	51	102
MTRBOAT					
NAT_{NAT}	−34	na	−48	130	71
NAT_{CR}	−46	na	−52	116	71
HIKE					
NAT_{NAT}	−23	−68	−4	71	42
NAT_{CR}	−37	−77	−11	61	46
BIKE					
NAT_{NAT}	−39	−61	−62	na	54
NAT_{CR}	−56	−79	−66	na	67
DHSKI					
NAT_{NAT}	na	na	54	66	60
NAT_{CR}	na	na	43	57	50
XSKI					
NAT_{NAT}	20	na	206	na	113
NAT_{CR}	−1	na	184	na	93
SNOWMOB					
NAT_{NAT}	na	na	−57	na	57
NAT_{CR}	na	na	−64	na	64
BGHUNT					
NAT_{NAT}	4	32	7	2	11
NAT_{CR}	−10	5	1	−3	5
SMHUNT					
NAT_{NAT}	−6	na	39	na	23
NAT_{CR}	−23	na	29	na	26
WATFOWL					
NAT_{NAT}	39	156	25	48	67
NAT_{CR}	20	100	18	41	43
FISH					
NAT_{NAT}	38	59	6	18	30
NAT_{CR}	18	24	0	13	14
WLVIEW					
NAT_{NAT}	27	19	9	−16	18
NAT_{CR}	5	−14	1	−21	10
ROCKCL					
NAT_{NAT}	−13	na	80	na	47
NAT_{CR}	−20	na	74	na	47
GENREC					
NAT_{NAT}	126	140	12	119	99
NAT_{CR}	87	73	4	106	67
OTHERREC					
NAT_{NAT}	na	37	−32	na	33
NAT_{CR}	na	−3	−36	na	20

Results are given in percent; na, not applicable, no benchmark estimates for activity in region.

Additionally, the in-sample convergent validity test results are consistent with percent difference findings for other types of convergent validity tests of benefit transfers [*Loomis*, 1992; *Smith and Huang*, 1995; *Loomis et al.*, 1995; *Downing and Ozuna*, 1996; *Kirchhoff et al.*, 1997; *Desvousges et al.*, 1998; *Kirchhoff*, 1998].

A closer examination of the performance of the NAT model in the two treatments is provided via the microlevel statistics reported in Table 6. The NAT model forecasted within 50% of the benchmark values primarily for those activities with a relatively large number of observations (BGHUNT, FISH, and WLVIEW). Conversely, where the NAT model forecasted values in excess of 100% difference from the benchmark values, the activities had relatively few observations (SWIM, NOMTRBT, MTRBOAT, XSKI, and GENREC). The notable exceptions to the above observations include PICNIC and SMHUNT for the former and WATFOWL for the latter.

To further investigate the relationship between the number of observations (n_{ij}) for each $\bar{y}_j^r$ (reported in Table 1) and the absolute value of the Δy_j^r reported in Table 6, we looked at two statistics. The correlation between n_{ij} and Δy_j^r is 0.29 for the $\mathrm{NAT_{NAT}}$ treatment and 0.41 for the $\mathrm{NAT_{CR}}$ treatment. The Δy_j^r are significantly and negatively related to the amount of information for each in both treatments (for $\mathrm{NAT_{NAT}}$, $\Delta y_j^r = 64.80 - 1.136 n_{ij}$, adjusted R^2 equals 0.07, F-statistic equals 5.26, degrees of freedom equal 58, and p value equals 0.026; and for $\mathrm{NAT_{CR}}$, $\Delta y_j^r = 57.81 - 1.210 n_{ij}$, adjusted R^2 equals 0.16, F-statistic equals 11.67, degrees of freedom equal 58, and p value equals 0.001). This means that as the number of observations increases for each activity per region, the percent difference between the forecasted and benchmark values decreases. That is, the more information the benefit transfer meta-analysis is based on at the activity per region level, the greater the precision of the model in forecasting the value.

4. National Model Forecast Values

Table 7 reports the forecast values from applying the national benefit transfer meta-analysis model using both treatments ($\mathrm{NAT_{NAT}}$ and $\mathrm{NAT_{CR}}$) for outdoor recreation activities. This application of the national model is identical to that used in testing for in-sample convergent validity. That is, each variable in the national model is held constant at its national mean value (reported in Table 3) or Census Region mean value with the exception of the activity variables when relevant. For all activities that do not have an activity dummy variable in the model, the forecast value is used for an activity composite, thus explaining the invariance in the forecast value for swimming, boating, off-road driving, hiking/backpacking, skiing, horseback riding, etc. Only those activities that have a representative activity dummy variable in the model will show some variance in their forecasted values. Table 7 provides values for every cell (activity by region), thus enabling forecast values for filling in empty cells in Table 1.

It is evident from Table 7 that the $\mathrm{NAT_{CR}}$ treatment provides lower forecasts in every case than the $\mathrm{NAT_{NAT}}$ treatment. In addition, the $\mathrm{NAT_{CR}}$ treatment provides a little more variation in forecast values for an activity across regions that the $\mathrm{NAT_{NAT}}$ treatment, in part because of the former treatment's increased sensitivity to difference across regions compared to the $\mathrm{NAT_{NAT}}$ treatment. Some additional variation (and possibly precision) in the forecast values for all activities from the National model can be attained by adapting some of the other variables, such as LAKE and RIVER. For example, fishing values can be adapted according to whether it is fishing in a lake or a river. However, confidence in these values would require out-of-sample testing if the adaptations are outside of the historic bounds of the database.

Table 7. Forecast Average Values for Outdoor Recreation Activities Using the National Benefit Transfer Meta-analysis Model

Activity/Value	CR1	CR2	CR3	CR45
Water-Related				
Swimming, float boating, and motorboating				
$\mathrm{NAT_{NAT}}$	34.46	35.17	35.82	34.80
$\mathrm{NAT_{CR}}$	28.40	25.30	33.24	32.76
Waterfowl hunting				
$\mathrm{NAT_{NAT}}$	44.62	45.33	45.98	44.96
$\mathrm{NAT_{CR}}$	38.56	35.46	43.40	42.92
Fishing				
$\mathrm{NAT_{NAT}}$	43.51	44.23	44.88	43.86
$\mathrm{NAT_{CR}}$	37.46	34.36	42.30	41.82
Non-Water-Related				
Camping, picnicking, sightseeing, hiking/backpacking, downhill and cross country skiing, small game hunting, wildlife viewing, and general and other recreation				
$\mathrm{NAT_{NAT}}$	34.46	35.17	35.82	34.80
$\mathrm{NAT_{CR}}$	28.40	25.30	33.24	32.76
Biking				
$\mathrm{NAT_{NAT}}$	20.89	21.60	22.25	21.23
$\mathrm{NAT_{CR}}$	14.83	11.74	19.67	19.19
Snowmobiling				
$\mathrm{NAT_{NAT}}$	14.16	14.87	15.52	14.50
$\mathrm{NAT_{CR}}$	8.10	5.00	12.94	12.46
Big game hunting				
$\mathrm{NAT_{NAT}}$	46.93	47.65	48.30	47.28
$\mathrm{NAT_{CR}}$	40.88	37.78	45.72	45.24
Off-road driving				
$\mathrm{NAT_{NAT}}$	26.56	27.27	27.92	26.90
$\mathrm{NAT_{CR}}$	20.50	17.41	25.34	24.86
Rock climbing				
$\mathrm{NAT_{NAT}}$	74.19	74.91	75.56	74.54
$\mathrm{NAT_{CR}}$	68.14	65.04	72.98	72.50
Horseback riding				
$\mathrm{NAT_{NAT}}$	22.62	23.33	23.98	22.96
$\mathrm{NAT_{CR}}$	16.56	13.46	21.40	20.92

$\mathrm{NAT_{NAT}}$ applies the national benefit transfer meta-analysis model holding explanatory variables except activity ones at their national mean values (Table 3). $\mathrm{NAT_{CR}}$ applies the national benefit transfer meta-analysis model holding explanatory variables except activity ones at their Census Region mean values. Values are given in U.S. dollars (1996 fourth quarter).

5. Summary and Recommendations

We estimated national and Census Region meta-analysis models of an outdoor recreation use value database spanning studies from 1967 to 1998. We then tested the convergent validity of the models as benefit transfer tools. We found that all of the models fit the data reasonably well. However, the national model minimized percent difference across all activities by region. The national meta-analysis regression model would be a good starting point for agencies estimating recreation benefits arising from broad environmental policies that effect recreation use over a large number of sites in a region. Such situations include non-point-source water quality improvements along entire river courses or major watersheds. Of course, the user should evaluate any estimate from a meta-analysis for reasonableness. Meta-analysis will never likely be a panacea for benefit transfers.

While meta-analysis may be a conceptually sound approach to benefit transfer, the quality of original research and full reporting of data and results is as necessary a component to critical meta-analysis as the statistical methods used. A meta-analysis can be no better than the data that it is built on. This limitation is exacerbated when the intent of original research is taken into account. Original research is most often not designed for future benefit transfer applications. In order for benefit transfer to become a secondary function of original research, guidelines for conducting original research including consistent data collection and reporting seem to be necessary. However, at the same time, we do not want to slow down investigations on the frontier of recreation research.

The ability of meta-analysis to capture nuances in the data, differences between sites, user populations, and/or affected activities, is dependent not only upon the quality of the original studies but also on the sheer volume of studies conducted. One of the limitations of our meta-analysis is the lack of an adequate number of studies for certain recreation activities. We provide evidence that as the amount of information increases, the precision of the meta-analysis forecasting increases. Therefore separate meta-analyses of different recreation activities, given enough observations, may provide models that are more robust to factors affecting them and therefore have an increased ability to function for benefit transfer compared to the all-encompassing model we presented and tested. This also means that more research on similar activities should be conducted under different conditions, in different regions, and using different methods.

As *Desvousges et al.* [1998] remind us, an important component in any benefit transfer is the involvement and judgment of the transfer analyst. The above limitations identified with using meta-analysis for benefit transfer are costs associated with the method, which should be weighed against the costs and benefits of primary research. We did not make any qualitative judgments concerning the exclusion of individual studies from the meta-analysis. Instead, all studies that could be fully coded were included in the analysis. Our approach may have led to a small numbers problem of identifying spurious correlation as systematic effects in the meta-analysis model. That is, a large or small value based on a few estimates for an activity in a region may be in the tails of the distribution around the true mean value for that activity. This "outlier" can have significant effects on the specification of the model. Therefore any meta-analysis is temporally relative and should not be treated as absolute.

The transfer specialist's confidence in the use of meta-analysis for benefit transfers may be increased with additional information regarding out-of-sample convergent validity testing of the application. These tests would better mirror real application contexts than in-sample testing. However, out-of-sample testing requires a sufficient number of original studies not included in the existing database. Therefore out-of-sample testing is an issue for future research as new studies on outdoor recreation use values are conducted.

Acknowledgments. This study was funded, in part, by the USDA Forest Service, Research Joint Venture 28-JV7-962. We thank Linda Langner, Ross Arnold, and Dan McCollum for their insights and Ram Shrestha for assisting with the data. This manuscript greatly benefited from the comments of Molly Espey and another referee of this journal and comments from participants at the USDA Regional Research Project W-133 meetings in Tucson, Arizona, 1999. Any errors or shortcomings one may find with this analysis are the sole responsibility of the authors.

References

Boyle, K., G. Poe, and J. Bergstrom, What do we know about groundwater values?: Preliminary implications from a meta analysis of contingent-valuation studies, *Am. J. Agric. Econ.*, *76*(5), 1055–1061, 1994.

Brookshire, D. S., and H. R. Neill, Benefit transfers: Conceptual and empirical issues, *Water Resour. Res.*, *28*(3), 651–655, 1992.

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

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.

Downing, M., and T. Ozuna Jr., Testing the reliability of the benefit function transfer approach, *J. Environ. Econ. Manage.*, *30*(3), 316–322, 1996.

Espey, M., J. Espey, and W. D. Shaw, Price elasticity of residential demand for water: A meta-analysis, *Water Resour. Res.*, *33*(6), 1369–1374, 1997.

Kirchhoff, S., Benefit function transfer vs. meta-analysis as policy making tools: A comparison, paper presented at the Workshop on Meta Analysis and Benefit Transfer, Tinbergen Inst., Amsterdam, April 6–7, 1998.

Kirchhoff, S., B. G. Colby, and J. R. LaFrance, Evaluating the performance of benefit transfer: An empirical inquiry, *J. Environ. Econ. Manage.*, *33*(1), 75–93, 1997.

Loomis, J. B., The evolution of a more rigorous approach to benefit transfer: Benefit function transfer, *Water Resour. Res.*, *28*(3), 701–705, 1992.

Loomis, J., and D. White, Economic benefits of rare and endangered species: Summary and meta-analysis, *Ecol. Econ.*, *18*(3), 197–206, 1996.

Loomis, J., B. Roach, F. Ward, and R. Ready, Testing the transferability of recreation demand models across regions: A study of Corps of Engineers reservoirs, *Water Resour. Res.*, *31*(3), 721–730, 1995.

MacNair, D., *1993 RPA Recreation Values Database*, Resour. Plann. Act Program, U.S. Dep. of Agric. For. Serv., Washington, D. C., 1993.

Rosenthal, D., The necessity of substitute prices in recreation demand analyses, *Am. J. Agric. Econ.*, *69*(4), 828–837, 1987.

Smith, V. K., and J. Huang, Can markets value air quality?: A meta-analysis of hedonic property value models, *J. Polit. Econ.*, *103*(1), 209–227, 1995.

Smith, V. K., and Y. Kaoru, Signals or noise?: Explaining the variation in recreation benefit estimates, *Am. J. Agric. Econ.*, *72*(2), 419–433, 1990.

Smith, V. K., and L. Osborne, Do contingent valuation estimates pass a scope test?: A meta-analysis, *J. Environ. Econ. Manage.*, *31*(3), 287–301, 1996.

Stanley, T. D., and S. B. Jarrell, Meta-regression analysis: A quantitative method of literature surveys, *J. Econ. Surv.*, *3*(2), 161–170, 1989.

U.S. Forest Service, *Draft 1990 RPA Program*, U.S. Dep. of Agric., Washington, D. C., 1989.

U.S. Water Resources Council, Principles, standards, and procedures for water and related land resource planning, *Fed. Regist.*, *38*(174), part III, 24,778–24,945, 1973.

U.S. Water Resources Council, Procedures for evaluation of national economic development (NED) benefits and costs in water resources planning (Level C), *Fed. Regist.*, *44*(243), 72,892–72,976, 1979.

U.S. Water Resources Council, *Economic and Environmental Principles and Guidelines for Water and Related Land Resources Implementation Studies*, U.S. Govt. Print. Off., Washington, D. C., 1983.

Walsh, R. G., D. M. Johnson, and J. R. McKean, Review of outdoor recreation economic demand studies with nonmarket benefit estimates, 1968–1988, *Tech. Rep. 54*, Colo. State Univ., Fort Collins, 1988.

Walsh, R. G., D. M. Johnson, and J. R. McKean, Issues in nonmarket valuation and policy application: A retrospective glance, *West. J. Agric. Econ.*, *14*(1), 178–188, 1989.

Walsh, R. G., D. M. Johnson, and J. R. McKean, Benefit transfer of outdoor recreation demand studies: 1968–1988, *Water Resour. Res.*, *28*(3), 707–713, 1992.

J. B. Loomis and R. S. Rosenberger, Department of Agricultural and Resource Economics, Colorado State University, Fort Collins, CO 80523. (rrosenbe@lamar.colostate.edu)

(Received October 8, 1999; revised January 13, 2000;
accepted January 14, 2000.)