Unified Dataset for All Choice Decisions in Recreational Water Quality Survey:

Addendum to Economics of Environmental Improvement

By Jason Bell
Duke University
Durham, North Carolina

August 7, 2009

Prepared for
National Center for Environmental Economics
Office of Policy, Economics, and Innovation
US Environmental Protection Agency
Washington, DC 20460

Disclaimer

Although prepared with partial EPA funding, this report has neither been reviewed nor approved by the U.S. Environmental Protection Agency for publication as an EPA report. The contents do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Preface

This report serves as an addendum to the EPA report "Economics of Environmental Improvement," by W. Kip Viscusi (Vanderbilt University), Joel Huber (Duke University), and Jason Bell (Duke University), which can be found on the EPA website (http://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0496-01.pdf/$file/EE-0496-01.pdf).

The original report estimated values, tradeoffs, and demographic relationships for several aspects of water quality from answers elicited from a 25-minute survey of over 4,000 respondents. Each of these several aspects was examined separately, from different questions or sets of questions in that survey.

This report confirms those results and expands them, using a dataset that pools all of those questions and sets of questions into a single analysis that was suggested by an external reviewer of the project report.

This broader analysis confirms previous results with findings that are similar to those estimated in the main report, while accounting for all respondent choices.

Expanded results are made possible through the ability to directly compare aspects of water quality valuation that were considered by respondents across multiple question sets. Among the new results is a comparison of the relative valuation of fishing, swimming, and the aquatic environment attributes at the regional and national level. The emphasis people place on these components differs depending on the scope of the water quality improvement. Improvements to the aquatic environment, which are non-use values, are most highly valued from a national perspective, while from a regional perspective the use value of fishing is most highly valued.

W. Kip Viscusi
Vanderbilt Law School
Nashville, TN

Executive Summary

This memo discusses analysis of data from a recreational water quality survey fielded between November 2002 and October 2004. EPA Cooperative Agreement CR823604 and Grant R827423 funded that survey and subsequent report, Economics of Environmental Improvement.[1]

The EPA report at the end of the grant covering the recreational water quality survey, as well as published articles considered each question set separately. Estimates were made for:

- Regional water quality value,
- National water quality value,
- Discounting of delayed policies,
- Lake vs. river preference, and
- Fishing vs. swimming vs. aquatic environment preference.

The goal of this project was to combine the several different choice sets presented in the recreational water quality survey into a single model. Doing so enables features to be directly compared and allows for estimations of each survey feature while taking all the other choices into account.

Expanded new results were produced in areas where multiple factors were examined at the same time. These results included:

- The relative preference of fishing, swimming, and the aquatic environment differ based upon whether the policies affected national or regional water quality. For all questions, the distribution of preferences within water quality components reflects a preference for the aquatic environment (44.6%) over fishing (34.6%) with swimming (20.8%) having the lowest priority. However, the preference for aquatic environment is more pronounced when the improvements are national in scope, while fishing has the greatest importance for regional improvements.
- The increased value of national improvements over regional improvements was confirmed. National improvements are about 25% higher than an equal improvement that only affects the respondent's home region.
- Most primary conclusions estimated in analysis of the individual question sets are supported when including all decisions made by respondents in the survey.

Overall, the results of this analysis that addressed topics from the previous project report were similar to those estimates.

[1] That report can be found at: http://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0496-01.pdf/$file/EE-0496-01.pdf

- Regional water quality value has a median value of $15.82.
- National water quality value has a median value of $19.73.
- Lake improvements are valued more highly than river improvements, though this preference is not extreme (52% preference for lakes, 48% for rivers).
- Water quality features varied by the demographic characteristics of respondents in expected ways, most commonly with income, education, age, and environmental organization membership; how often respondents visited lakes and rivers was also a factor in explaining many of these variations.

Table of Contents

Preface…… 2

Executive Summary…… 3

Table of Contents..... 5

Introduction..... 6

Choice Models for Cost, Quality, and Delay..... 9

National vs. Regional..... 13

Hierarchical Bayes Results 15

Cost Utility..... 16

Improvement Utility..... 17

Cost – Improvement Tradeoff..... 18

Delay Utility..... 19

Lake and River Utility 20

Conclusions..... 22

Introduction

During the recreational water quality survey, 4,241 respondents were presented with 142,758 choices in 64,094 questions. Between two and three options were presented per question, and there were between one and 22 questions completed by respondents. The average respondent completed about 15 questions, with the median respondent completing 18.

The choices presented differed on some of the following features:

1. Cost of the choice
2. Water quality improvement of the choice
3. Components of water improved: fishing, swimming, and the aquatic environment
4. Delay before the choice began
5. Lakes improved by the choice
6. Rivers improved by the choice
7. Whether the choice improved regional or national water quality

All of these choices were combined into a single dataset for two sets of analyses.

The first analysis investigated choice decisions using a conditional logit model grouping choices into their associated questions, and clustered by individual respondent. This analysis estimated how the features of the questions determine which choices respondents selected.

The second analysis used Hierarchical Bayes to estimate the utility values of the features for each individual respondent. This analysis allowed an investigation of how individual respondent attributes are related to the importance respondents placed upon each feature.

The table below shows the mean value of each of the choice features, as well as the four demographic characteristics examined in the conditional logit models.

Variable	N	Mean	Standard Deviation	Min.	Max.
Cost	142758	$ 147.31	164.2901	0	600
Quality Improvement	142758	7.64 %	9.13528	-5	40
Fishing Quality	142758	7.66 %	9.31758	-5	40
Swimming Quality	142758	7.63 %	9.330435	-5	40
Aquatic Environment Quality	142758	7.62 %	9.34642	-5	40
Delay	142758	.992 years	1.898001	0	6
Lake	142758	.8730719	.3328936	0	1
River	142758	.8730719	.3328936	0	1
Nation	142758	.5188921	.4996447	0	1
Log(income)	139800	10.52623	.9097489	7.82	12.07
Years of Education	139800	13.17 years	2.561943	8	21
Age	139800	44.64 years	16.98609	18	98
Member of Environmental Org.	139800	.0556009	.2291501	0	1

The Cost value is simply the annual increase in cost of living associated with the choice. In the few questions that did not present a cost with the choice, the value was $0.

The similarity between the Quality Improvement value and the Fishing, Swimming, and Aquatic Environment values are due to the fact that Quality Improvement is an average of the three components. For questions that simply asked about water quality, those components were all equal, but for the questions that asked about individual components, they were different.

The three improvement components correspond to the definitions in the National Water Quality Inventory.[2] The definitions provided to respondents were:

Fishing: A lake or river is good for fishing if eating fish caught in the lake or river will not make you sick. A lake or river is not good for fishing if eating fish caught in the lake or river could make you sick.

Swimming: A lake or river is good for swimming if prolonged contact with the water in the lake or river will not make you sick. A lake or river is not good for swimming if prolonged contact with the water can make you sick.

Aquatic Environment: The aquatic environment is good if the lake or river supports a healthy, balanced community of aquatic life, such as fish, plants, insects, and algae.

The Delay value is the amount of time, in years, before the choice takes place. For most choices, changes would take place in the same year, so the value was 0. For other choices, delays ranged between 2 and 6 years.

The Lake and River values reflect whether the choice affects lakes or rivers. For most choices, both lakes and rivers were affected, so both of these values would be 1. For some choices, only lakes or only rivers were affected.

The Nation value describes whether the choice affected the entire country. If not, then the choice only affected the respondent's region (within 100 miles of the respondent's home, a 2-hour drive or so).

Respondents answered several types of questions in the survey.

The first type of question was an iterated choice set, where answers to earlier questions determined what question would be asked next. These questions included two choices (two regions for the region choice set, and a yes / no referendum for the national set) and a no preference option. If the respondent chose no preference, that question was excluded from analysis. Iterated question sets were asked for regional then national water quality improvement of combined components (fishing, swimming, and aquatic environment), of combined lakes and rivers, had no delay, and included a cost feature).

[2] The National Water Quality Inventory Report to Congress homepage can be found at: http://www.epa.gov/305b/

The second type of question was a conjoint question set, where multiple questions were asked, each of which took the same form, but the answer to any of them was unrelated to the question asked before or after. These questions included 2 to 3 choices where 3 features changed. In one conjoint set, the three factors were the change in quality of fishing, swimming, and the aquatic environment, where one choice was associated with lakes, and the other with rivers. In the other conjoint set, the three factors were cost, quality improvement, and delay. Conjoint question sets were asked either in a regional context or a national context.

While no question considered variations in every feature, and no two question sets considered exactly the same features, each question set had at least one feature in common with any other question set. This relationship across all questions is the basis for this analysis.

Choice Models for Cost, Quality, and Delay

The most simple conditional logit model predicts a respondent's choice based on just the cost, quality, and delay associated with a choice. As expected, larger cost and longer delays are less likely to be chosen and higher quality is more likely to be chosen. The trade-off in this model implies a willingness to exchange $17.49 for a 1 percentage point change in water quality, a value that roughly corresponds to the median value among respondents.

Conditional (fixed-effects) logistic regression

Number of obs = 138936
Wald chi2(3) = 2169.82
Prob > chi2 = 0.0000
Log pseudolikelihood = -47606.548
Pseudo R2 = 0.0286

(Std. Err. adjusted for 4220 clusters)

	Coefficient	Standard Error
Cost	-.0019013 ***	.0000683
Delay	-.1961558 ***	.0056657
Quality	.0331763 ***	.00096

This model was repeated, but with the quality feature replaced with its component parts: fishing quality, swimming quality, and aquatic environment quality. This model demonstrates that respondents tend to weight the aquatic environment feature (44.6% share) more highly than the fishing feature (34.6% share), and the swimming feature least of the three (20.8% share).

Conditional (fixed-effects) logistic regression

Number of obs = 138936
Wald chi2(5) = 2180.00
Prob > chi2 = 0.0000
Log pseudolikelihood = -47603.063
Pseudo R2 = 0.0287

(Std. Err. adjusted for 4220 clusters)

	Coefficient	Standard Error		Implied Preference Share of Quality Feature
Cost	-.0019024 ***	.0000682		
Delay	-.1962059 ***	.0056657		
Fishing	.0114964 ***	.0020994		34.6 %
Swimming	.0069103 ***	.0020403		20.8 %

Aquatic Environment	.0147927 ***	.0020476		44.6 %

The next model adds interaction terms for whether the choice was national in scope and interactions with demographic characteristics.

Respondents were less likely to choose higher cost choices, more likely to choose higher quality choices, and less likely to choose delayed choices if they were national.

As for demographics, higher income respondents were more likely to choose higher priced and higher quality options than lower income respondents. This is not surprising, as wealthier respondents can afford more of a good, even at higher prices than others.

More educated respondents were more likely to choose higher quality options. This result is most likely associated with also choosing higher cost options, though that is not a significant coefficient in the model.

Older respondents were less likely to choose options with longer delays, which could be because older people do not have as much time to enjoy water quality amenities, or perhaps older respondents were more likely to believe that a delay in a policy might ultimately result in the policy not being done at all.

Finally, environmental organization members were more likely to choose higher cost options than other respondents, likely related to the correlation between higher cost options and higher quality options in many of the survey questions.

Conditional (fixed-effects) logistic regression
Number of obs = 136012
Wald chi2(18) = 2301.84
Prob > chi2 = 0.0000
Log pseudolikelihood = -46227.725
Pseudo R2 = 0.0368

(Std. Err. adjusted for 4015 clusters)

	Coefficient	Standard Error
Cost	-.0020907 ***	.0000701
Quality	.037099 ***	.0010379
Delay	-.2057421 ***	.0058625
Cost X National	-.0005066 ***	.0001216
Quality X National	.0163378 ***	.0019363
Delay X National	-.0461758 ***	.0115779
Cost X Log(income)	.000308 ***	.0000839
Cost X Years of education	.0000233	.0000277
Cost X Age	.0000191 ***	4.02e-06
Cost X Environmental org. member	.0013868 ***	.0002592

Quality X Log(income)	.0028693 ***	.0010962
Quality X Years of education	.0025154 ***	.0004034
Quality X Age	.0000784	.0000578
Quality X Environmental org. member	.0057961	.0053327
Delay X Log(income)	.00272	.0061454
Delay X Years of education	-.0019276	.0023131
Delay X Age	-.0023555 ***	.0003512
Delay X Environmental org. member	-.0292981	.0259808

National vs. Regional

In the water quality survey, respondents were not offered questions where they could choose discretely between regional and national choices. Some questions involved a set of choices that were national in scope, and other questions involved a set of choices that only affected the respondent's region. For this reason, it was not possible to construct a logit model with national as a variable, though it could be interacted with other features.

However, parallel models can be estimated on the two sets of choices (national and regional), and the tradeoffs can be compared to determine the relative preference between regional and national choices. These models indicate that national improvements are more valuable than regional improvements to respondents, by almost 25%.

Region Choices		Nation Choices
Conditional (fixed-effects) logistic regression Number of obs = 66978 Wald chi2(3) = 930.55 Prob > chi2 = 0.0000 Log pseudolikelihood = -23040.187 Pseudo R2 = 0.0246 (Std. Err. adjusted for 3926 clusters)		Conditional (fixed-effects) logistic regression Number of obs = 71958 Wald chi2(3) = 1244.67 Prob > chi2 = 0.0000 Log pseudolikelihood = -24518.435 Pseudo R2 = 0.0342 (Std. Err. adjusted for 3743 clusters)

	Region Choices			Nation Choices	
	Coefficient	Standard Error		Coefficient	Standard Error
Cost	-.0017052 ***	.0000876		-.0021751 ***	.000094
Quality	.0269724 ***	.0011665		.0429176 ***	.0015548
Delay	-.1772122 ***	.0079585		-.2221636 ***	.0082334
Implied Tradeoff					
$ per 1% point	$15.82			$19.73	

Parallel models can also estimate how the relative importance of individual components of water quality (fishing, swimming, and aquatic environment) differs between regional and national choices.

In choices affecting only the respondent's home region, fishing (52%) is considered most important, followed by the aquatic environment (30%), followed by swimming (18%).

In choices that affect the entire country, the aquatic environment (47%) is considered most important, followed by fishing (29%), followed by swimming (23%).

These preference differences are understandable, with the aquatic environment (the feature most closely associated with non-use existence value) valued more highly for improvements where the respondent is less able to directly make use of it. The two features with direct use value become closer to each other in value for national improvements.

Region Choices		Nation Choices
Conditional (fixed-effects) logistic regression Number of obs = 66978 Wald chi2(5) = 933.16 Prob > chi2 = 0.0000 Log pseudolikelihood = -23038.655 Pseudo R2 = 0.0247 (Std. Err. adjusted for 3926 clusters)		Conditional (fixed-effects) logistic regression Number of obs = 71958 Wald chi2(5) = 1255.63 Prob > chi2 = 0.0000 Log pseudolikelihood = -24514.241 Pseudo R2 = 0.0344 (Std. Err. adjusted for 3743 clusters)

	Region Choices			Nation Choices	
	Coefficient	Standard Error		Coefficient	Standard Error
Cost	-.0017056 ***	.0000876		-.0021768 ***	.000094
Delay	-.1772318 ***	.0079576		-.2222545 ***	.0082341
Fishing	.0141574 ***	.0035631		.01258 ***	.0025919
Swimming	.0047713	.0034541		.0100821 ***	.0025545
Aquatic Environment	.0080527 **	.0034412		.0202969 ***	.0025506

Hierarchical Bayes Results

A Hierarchical Bayes analysis was performed to estimate individual utilities for the cost, improvement, delay, lake, river, national, and base quality features in the choices.

A positive utility for a feature means that respondents considered a higher value in that feature more desirable. A negative utility for a feature means that a higher value in that feature was considered less desirable.

As expected, cost and delay features had negative utilities, while improvement, lake, and river features had positive utilities. In addition, national effects have a positive utility, meaning they are preferred over choices only affecting the respondent's region. Finally, choices with higher starting quality were preferred to choices with lower starting quality.

Variable	Obs	Mean	Standard Deviation	Minimum	Maximum
Cost Utility ($1)	4220	-.0047986	.003511	-.01347	.00365
Improvement Utility (1%)	4220	.0725674	.0384134	-.0344	.1725
Delay Utility (1 year)	4220	-.3589066	.2575581	-1.055	.442
Lake Utility	4220	1.713304	.1780775	1.095	2.299
River Utility	4220	1.597053	.1759169	.998	2.183
National Utility	4220	.0367773	.1065166	-.347	.421
Base quality Utility (1%)	4220	.1460324	.0111547	.1028	.1882

While there is no unit associated with these utilities, they are standardized across features and respondents. That means they can reveal differences in preference between two respondents, as well as a single respondent's tradeoff between features.

For instance, a respondent with average utility values would trade off improvement with cost at a rate of 0.073 / 0.0048, or $15.12 for a one percentage point change in water quality.

The utilities estimated by the Hierarchical Bayes analysis can also be examined by how utility for each feature differs on demographic characteristics of the respondent.

Cost Utility

For the cost utility, several demographic characteristics are associated with a greater utility for cost (or, less aversion to higher costs). Among these are higher income, more educated, older, environmentalist, trip taking (both in and out of region), and Hispanic respondents. Also, respondents with greater lake density in their home state had greater cost utility. Respondents in larger households had a greater aversion to greater costs.

Regression Model

Number of obs = 4015
F(16, 3998) = 14.78
Prob > F = 0.0000
R-squared = 0.0559
Adj R-squared = 0.0521

	Coefficient	Standard Error
Log(income)	.0003574 ***	.0000646
Years of education	.0000773 ***	.0000226
Age	.0000217 ***	3.42e-06
Environmental org. member	.0014024 ***	.0002476
Trips to lakes or rivers	.0000567 ***	.0000171
Trips outside region	.0002078 ***	.0000563
Race: black	.0000219	.0001681
Race: nonblack, nonwhite	3.72e-06	.0002307
Hispanic	.0004555 **	.0001848
Female	.000085	.0001102
Household size	-.0000979 **	.000043
Region: Northeast	.0001111	.000172
Region: South	.0001757	.0001565
Region: West	-.0000816	.0001709
State lake quality	6.81e-07	2.16e-06
Lake acres per square mile in state	.0000113 *	5.98e-06
Intercept	-.0110411 ***	.0007179

Improvement Utility

Respondents with higher utility for improvements (or, a relatively greater desire for improvements to lakes and rivers) have many of the same demographic characteristics as those with higher cost utility. Higher income, more educated, older and environmentalist respondents, as well as those who took more trips, and respondents who live in states with higher lake density had higher utility for improvements. Black respondents had a lower utility for such improvements compared to other respondents.

Number of obs = 4015
F(16, 3998) = 17.84
Prob > F = 0.0000
R-squared = 0.0666
Adj R-squared = 0.0629

	Coefficient	Standard Error
Log(income)	.0035922 ***	.0006999
Years of education	.0017751 ***	.0002444
Age	.0001227 ***	.000037
Environmental org. member	.0080906 ***	.0026829
Trips to lakes or rivers	.0008673 ***	.0001852
Trips outside region	.0022588 ***	.0006104
Race: black	-.0083434 ***	.0018212
Race: nonblack, nonwhite	-.0038231	.0025
Hispanic	.0004097	.0020028
Female	-.0010243	.0011937
Household size	-.0006836	.000466
Region: Northeast	.002581	.0018634
Region: South	.0003427	.0016952
Region: West	-.0009032	.0018522
State lake quality	9.69e-06	.0000234
Lake acres per square mile in state	.0001404 **	.0000648
Intercept	.0022946	.007778

Cost – Improvement Tradeoff

A ratio can be constructed to estimate each respondent's tradeoff between cost and improvement by dividing the cost utility into the improvement utility, as was done for the overall mean utilities in Table X.

Using individual respondent utilities leads to a higher mean tradeoff than overall means, $23.96.

The log of this ratio was used in the regression model below.

Respondents with a higher cost per unit improvement tend to have higher incomes, more years of education, are older, and take more trips to lakes and rivers both in and outside of their home region. Black respondents tended to have a lower cost per improvement value than other respondents.

Number of obs = 3648
F(16, 3631) = 12.09
Prob > F = 0.0000
R-squared = 0.0506
Adj R-squared = 0.0464

Log(income)	.1077402 ***	.0279602
Years of education	.0554143 ***	.0098354
Age	.0080819 ***	.0014821
Environmental org. member	.3224777 ***	.113932
Trips to lakes or rivers	.0307671 ***	.0073952
Trips outside region	.074835 ***	.0246519
Race: black	-.1233812 *	.0726266
Race: nonblack, nonwhite	.0107946	.1002903
Hispanic	.0493443	.0810277
Female	.0502381	.0477275
Household size	-.0257214	.0185058
Region: Northeast	.0338784	.07479
Region: South	.0297543	.0677434
Region: West	-.0371128	.0739317
State lake quality	.0005428	.0009324
Lake acres per square mile in state	.0040733	.0025928
Intercept	.3241151	.3097247

Delay Utility

Respondents with lower utility for delay have a relatively greater desire for choices that take place sooner. In previous published research[3], the discount rate for respondents in this survey was estimated at around 10%, depending upon the length of the delay.

A regression model of the delay utility shows that older respondents, environmentalists, those who take more trips to lakes and rivers outside their region, Black, and Hispanic respondents all have a greater preference for choices that occur sooner, relative to other respondents.

Number of obs = 4015
F(16, 3998) = 8.41
Prob > F = 0.0000
R-squared = 0.0326
Adj R-squared = 0.0287

	Coefficient	Standard Error
Log(income)	-.0077269	.0048002
Years of education	-.0003048	.0016763
Age	-.0020276 ***	.0002539
Environmental org. member	-.0314749 *	.0184004
Trips to lakes or rivers	-.0007423	.00127
Trips outside region	-.0106425 **	.0041866
Race: black	-.0546413 ***	.0124907
Race: nonblack, nonwhite	.012827	.0171457
Hispanic	-.0456119 ***	.0137358
Female	-.018202 **	.008187
Household size	.0030751	.0031963
Region: Northeast	.0128038	.0127797
Region: South	-.0129181	.011626
Region: West	.0135447	.0127028
State lake quality	.0000279	.0001604
Lake acres per square mile in state	-.0003573	.0004444
Intercept	-.1554357 ***	.0533446

[3] Viscusi, W. Kip, Joel Huber, and Jason Bell, "Estimating Discount Rates for Environmental Quality from Utility-Based Choice Experiments," Journal of Risk and Uncertainty, Vol. 37, No. 2/3 (December 2008), pp. 199-220.

Lake and River Utility

For most of the water quality survey, all choices affected both lakes and rivers. However, for one set of questions, respondents could choose between two choices, one of which affected only lakes, and the other affected only rivers (water quality for fishing, swimming, and the aquatic environment also varied in these questions). A logit model estimating choices including these features could not be constructed, but the Hierarchical Bayes analysis estimated utilities for these features.

The utilities themselves reflect a slight preference for lakes over rivers among respondents. This preference reflects a 52% / 48% tradeoff preference in favor of lakes.

Variable	Obs	Mean		Implied Percentage Share
Lake Utility	4220	1.713304		51.8 %
River Utility	4220	1.597053		48.2 %

The differences between these utilities in terms of demographic factors were interesting, and below they are compared side-to-side. Higher income respondents valued rivers more and lakes less. Those who took trips to lakes and rivers value rivers more. Hispanics and non-white non-black respondents value lakes more and rivers less. Female respondents value rivers less. Northeastern and Western respondents value lakes less and rivers more, and Southern respondents value rivers more.

The opposite movement of utilities for significant demographic variables is not surprising, since respondents had to choose one to the exclusion of the other in the questions involving lakes and rivers.

Lake Utility		River Utility
Number of obs = 4015 F(16, 3998) = 1.79 Prob > F = 0.0264 R-squared = 0.0071 Adj R-squared = 0.0032		Number of obs = 4015 F(16, 3998) = 3.96 Prob > F = 0.0000 R-squared = 0.0156 Adj R-squared = 0.0116

	Lake Utility			River Utility	
	Coefficient	Standard Error		Coefficient	Standard Error
Log(income)	-.0060229 *	.0032784		.0115034 ***	.0032083
Years of education	.0007953	.0011449		.0027746 **	.0011204
Age	.0003672 **	.0001734		.0000425	.0001697
Environmental org. member	-.0170002	.0125671		.0188155	.0122985
Trips to lakes or rivers	-.000775	.0008674		.0024677 ***	.0008488

Trips outside region	.0008605	.0028594		-.0018101	.0027982
Race: black	.0023155	.0085309		-.00873	.0083485
Race: nonblack, nonwhite	-.0224757 *	.0117102		.0208596 *	.0114598
Hispanic	-.0197523 **	.0093813		.015481 *	.0091807
Female	.0073667	.0055916		-.011256 **	.005472
Household size	.0022846	.002183		-.0031157	.0021364
Region: Northeast	-.0207313 **	.0087283		.0193578 **	.0085417
Region: South	-.0122419	.0079404		.018994 **	.0077706
Region: West	-.015569 *	.0086757		.0146165 *	.0084903
State lake quality	-.000068	.0001096		.0000421	.0001072
Lake acres per square mile in state	.0002685	.0003035		-.0003188	.000297
Intercept	1.758436 ***	.0364334		1.429221 ***	.0356545

Conclusions

This research demonstrates the advantage of combining the substantial number of different tasks performed by respondents in the recreational water quality survey. The resulting dataset contains over 140,000 choices considered by over 4,000 respondents, who made choices on the basis of differences in cost, water quality improvement, individual components of recreational water quality improved, delay in implementation of the choice, whether the choice improved lakes or rivers or both, and whether the choice was national in scope or affected only the respondent's home region.

Combining all respondent choices into a single analysis yielded results that reinforce those found in individual analyses. However, the expanded analysis also yielded insight into important differences in preference shares of individual water quality components (fishing, swimming, and aquatic environment). The pooled analysis also confirms a higher value for national improvements that was not clear from individual analysis.

Matching the demographic characteristics of the respondents with the choices they considered and preferred gave new insight into the basis of choices made. Income, education, age, and environmental organization membership were more often significant drivers of respondents' tradeoffs between cost, quality improvement, and delay associated with choices.

The pooled Hierarchical Bayes analysis allowed a more in-depth examination of these effects, estimating individual utilities for each choice feature. The utilities from the new analysis were more sensitive to demographic characteristics and how often respondents took advantage of recreational water quality amenities. Higher cost and improvement utilities, as well as the cost per unit improvement ratio are all related to higher income, more education, greater age, environmental organization membership, and more trips made to lakes and rivers among respondents. Lower delay utilities were related to greater age, environmental organization membership, more lake and river trips made outside the home region, black, Hispanic, and female respondents. Lake and river utilities were also sensitive to a variety of demographic characteristics, often in opposite directions (when lake utilities fell, river utilities rose). Those included income, education, age, trips to lakes and rivers, races other than black or white, Hispanic, female, and those in the Northeast and West of the United States.

What is particularly interesting about this analysis is the ability to examine relationships between features, since all features are considered at the same time in the dataset, even though respondents may not have considered them at the same time in any single question. For instance, individual water quality features (fishing, swimming, and aquatic environment) change preference order depending on national or regional context in ways that are consistent with a use value vs. non-use value relationship. Specifically, since respondents cannot use the vast majority of lakes or rivers in the country, they value the non-use existence of a healthy aquatic environment more than the other components when choices are national in scope. Conversely, since respondents can use a much greater portion of lakes and rivers in their home region, the use benefit of fishing rises to prominence in regional choices, though a healthy aquatic environment still retains some importance.

In summary, the pooled dataset permitted greater power in analysis, particularly in the consideration of interactions between preferences for components of water quality as well as the characteristics of the respondents making choices about water quality features.