MEMORANDUM September 14, 2015

TO Katherine Pease, NOAA

FROM Total Value Team

SUBJECT Technical Memo TM-10: Econometric Analysis of Choice Questions

10.1 INTRODUCTION

This technical memo has three objectives:

- To test the cost and quantity conditions associated with consistent economic choices using the survey respondents' votes for each set of injuries separately;
- To test the hypotheses conventionally associated with the National Oceanic and Atmospheric Administration (NOAA) Panel's scope test using the survey respondents' votes;
- To estimate linear probability and Lewbel-Watanabe multivariate models to evaluate the effects of economic variables that are conventionally hypothesized as determinants of the survey respondents' votes;

The Total Value (TV) contingent valuation survey was designed to estimate a lower bound for total loss in economic value from the injuries to Natural Resource Trustee resources due to the release of oil into the Gulf of Mexico from the Macondo well in 2010. Two sets of injuries were provided to the TV team: a smaller set, labeled Version A injuries and a larger set, labeled Version B.[1]

The memo has five sections, including this introductory section. Section 10.2 describes the data. After that, section 10.3 presents the results of tests of the cost condition defined in Technical Memo TM-3. The fourth section presents the results for two different methods for conducting a multivariate analysis of the determinants of the sample respondents' votes for the prevention program described in Technical Memo TM-3. The last section, 10.5, summarizes the results for the tests for scope. The first set of tests uses linear probability models and the second test uses a nonparametric method for estimating the mean willingness to pay implied by these votes.

[1] The specific details for the two sets of injuries are given in Technical Memo TM-2.

10.2 DATA

10.2.1 OVERVIEW OF THE SURVEY

As noted in Technical Memo TM-7, 3,965 interviews were completed. The validation process identified 309 interviews that were removed for reasons described in Technical Memo TM-7.[2] The remaining 3,656 interviews are described here as *the analysis sample*. The first of these interviews began on October 27, 2013 and the last on July 17, 2014.

Each survey respondent was randomly assigned one of the 10 different variants of the questionnaire. Two factors distinguish these variants: the version characterizing the injuries (A and B) and the five one-time tax amounts ($15, $65, $135, $265, $435). Table 1 summarizes the disposition of the final analysis sample across the two versions and five tax amounts: 1,833 observations were randomly assigned to Version A and 1,823 to Version B. Samples ranging between 716 and 747 respondents received one of the five tax amounts. As expected, the distribution of respondents assigned to each tax amount is about equally divided between the two versions.

TABLE 1. RANDOM ASSIGNMENT OF SAMPLE TO TAX AMOUNTS AND VERSIONS

Tax Amount	Version		
	A	B	Total
15	368	364	732
65	370	377	747
135	368	366	734
265	371	356	727
435	356	360	716
Total	1,833	1,823	3,656

All of the analysis associated with this assessment of respondents' votes as economic choices was conducted without the sampling weights. Our purpose here is not to determine the economic responses in the entire population, but rather to test whether the respondents to the survey made choices that are consistent with the conditions discussed in Technical Memo TM-3. Random assignment of the questionnaires that vary in the injury description and the tax amounts ensures that we can obtain valid estimates on the subpopulation represented by the respondents.[3]

[2] See Appendix 1.10 for further analysis of these interviews.

[3] While we do not use sampling weights in estimation, we do adjust the standard errors and test statistics to account for the stratified sampling and the clustering at the modified PSU level. Twenty-two PSUs were sampled with certainty. As a

The descriptions for the coding of all variables based on respondents' answers to the questionnaire are given in Technical Memo TM-13. Table 1 in that memo summarizes the variable names and the coding procedures for the responses.

10.3 TESTS OF THE COST AND QUANTITY CONDITIONS FOR CONSISTENT ECONOMIC CHOICES

Random assignment of the one-time tax amounts and the injury descriptions allows separate testing of the cost condition using each of the subsamples associated with the Version A and Version B questionnaires. With statistically equivalent sub-samples receiving each tax amount for each injury description, the cost condition implies that the average proportion of votes for the program will not increase as the tax amount increases. This property is sometimes described as the "weak monotonicity" condition.

Figure 1 plots the percentage of respondents voting for the prevention plan on the vertical axis and the tax amount on the horizontal axis. As can be seen from the figure, only at the lowest tax amount, $15, does a majority of respondents vote for the plan – whether the injury scenario is Version A or Version B. For the larger injury, 49 percent of respondents vote for the plan at the next tax amount, $65. As discussed in Technical Memo TM-3, economic theory does not indicate what individuals' votes should be. These choices are based on individual tastes, the tax amount presented, and each individual's circumstances.

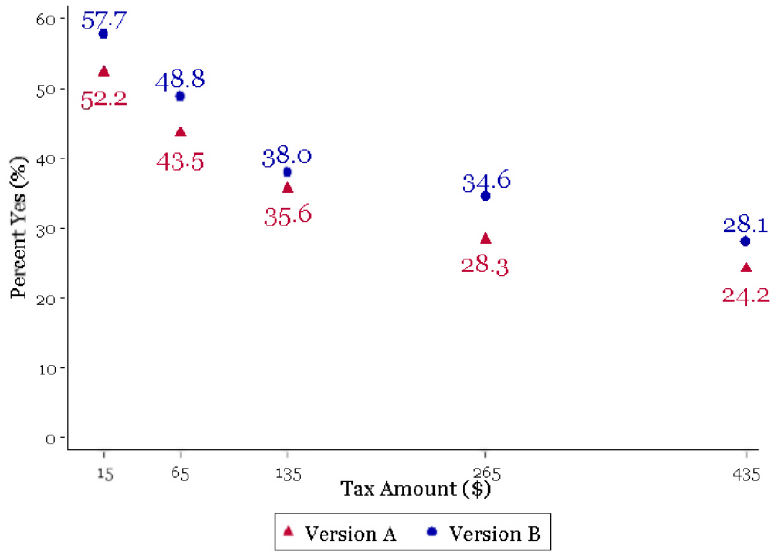
Figure 1 indicates that the proportion voting for the program declines with increases in the tax amount for each version of the injury description, demonstrating that the monotonicity condition. In fact, the stronger form of monotonicity holds in the sense that the yes vote percentages are strictly declining as the tax amount increases. Consequently, there is no evidence that weak monotonicity fails. In order to reject weak monotonicity, the proportions voting for the program would need to increase in moving from a smaller to a larger tax amount for at least one tax amount pair. Moreover, the increase in the proportion voting for the program would need to be statistically significant.

Figure 1 also illustrates that at each tax amount, the percentages voting for the prevention program with the Version A injury description lie below those for the larger injury description associated with the Version B.[4] This result indicates that there is no evidence against the quantity condition. More stringent statistical tests of the quantity condition are presented in section 10.5, along with the other tests for the NOAA scope condition

result, the SSUs within these twenty two PSUs were treated as the equivalent of PSUs for the purposes of clustering adjustments to the standard errors.

[4] The Version B description of injuries can be considered "larger" than Version A because the injuries described in Version A are a proper subset of those in Version B.

FIGURE 1.



It is possible to provide stronger statements about the monotonicity condition by determining whether the drop in the percentages voting yes from a smaller to larger cost value is statistically significant. The first step in this process tests the null hypothesis that the yes vote percentages are the same at each of the five tax amounts against the two-sided alternative that at least one pair is different. For each of the two injury scenarios this hypothesis is rejected using the Rao-Scott Corrected F-statistic (which accounts for Primary Sampling Unit (PSU) clustering as well as arbitrary heteroskedasticity). For Version A, F = 20.63 (p-value = 0.000) and for Version B, F = 22.41 (p-value=0.000). Thus, even using a two-sided alternative the null hypothesis that the percentage of yes votes does not vary with the tax amount is rejected.

Further support is provided by testing whether each drop in the yes vote percentage shown in Figure 1 is statistically significant. In other words, assume the hypothesis of monotonicity fails. This test considers whether there is a statistically significant rejection of non-monotonic responses. A statistical rejection of this null hypothesis is possible only if the data are consistent with strong monotonicity. This is a necessary, not a sufficient condition.

To determine whether the decline in the proportion of yes votes is statistically significant, this analysis considers the yes percentages at each pair of adjacent tax amounts: $15 to $65, $65 to $135, and so forth. The null hypothesis is that the percentage of yes votes is the same, and the one-sided alternative is that the vote percentage is smaller for the larger

of the two tax amounts. The test is performed using t statistics robust to clustering at the PSU level for the test of each univariate hypothesis[5].

Tables 4 and 5 contain the results of the t tests. For Version A, the null hypothesis of equality of the yes percentages is rejected in all four cases with one-sided p-values below 0.10. Three of the four tests reject with p-values below 0.05. The same tests for Version B reject the null hypothesis of equality of yes percentages at below the 0.05 level with one sided alternative hypotheses for three of the four comparisons. Only the comparison of $135 to $265 fails to reject the null hypothesis of equality at the 0.05 level (p-value = .158).

TABLE 2. VOTE TABULATED BY THE ONE-TIME TAX AMOUNT FOR THE VERSION A INJURY DESCRIPTION

	Vote				
	Against		For		Total
Tax Amount	N	Row %	N	Row %	N
15	176	48%	192	52%	368
65	209	57%	161	44%	370
135	237	64%	131	36%	368
265	266	72%	105	28%	371
435	270	76%	86	24%	356
Total	1,158	63%	675	37%	1,833

Rao-Scott Corrected F **20.627**
P-value 0.000

[5] Other statistics, such as a likelihood ratio statistic, would not be robust to cluster correlation, and is not as easily used for a one-sided alternative.

TABLE 3. VOTE BY THE ONE-TIME TAX AMOUNT FOR THE VERSION B INJURY DESCRIPTION

Tax Amount	Vote				
	Against		For		Total
	N	Row %	N	Row %	N
15	154	42%	210	58%	364
65	193	51%	184	49%	377
135	227	62%	139	38%	366
265	233	65%	123	35%	356
435	259	72%	101	28%	360
Total	1,066	59%	757	42%	1,823

Rao-Scott Corrected F **22.408**
P-value 0.000

TABLE 4. PAIR-WISE TESTS OF PROPORTIONS VOTING FOR PROGRAM FOR VERSION A

Version	Tax Amount 1	Tax Amount 2	Percent Yes 1	Percent Yes 2	Difference	t	One-sided P-value
A	15	65	52.2%	43.5%	-8.7%	-2.53	0.006
	65	135	43.5%	35.6%	-7.9%	-2.30	0.011
	135	265	35.6%	28.3%	-7.3%	-2.07	0.020
	265	435	28.3%	24.2%	-4.1%	-1.31	0.096

TABLE 5. PAIR-WISE TESTS OF PROPORTIONS VOTING FOR PROGRAM FOR VERSION B

Version	Tax Amount 1	Tax Amount 2	Percent Yes 1	Percent Yes 2	Difference	t	One-sided P-value
B	15	65	57.7%	48.8%	-8.9%	-2.27	0.012
	65	135	48.8%	38.0%	-10.8%	-3.18	0.001
	135	265	38.0%	34.6%	-3.4%	-1.00	0.158
	265	435	34.6%	28.1%	-6.5%	-1.91	0.029

10.4 MULTIVARIATE ANALYSIS OF VOTES

10.4.1 INTRODUCTION

Votes are discrete outcomes – for or against the prevention program. The multivariate analyses reported here use two different estimators. The first is the linear probability specification.[6] Ordinary least squares regression using a binary coding of votes (for = 1,

[6] Appendix 6.1 reports probit estimates. The overall conclusions were not affected by the estimates used for the multivariate analysis.

against = 0) allows the estimated effect of each independent variable to be interpreted as a measure of the change in the probability of a vote for the prevention program with a unit change in that variable.[7] With a completely specified willingness to pay (WTP) distribution, one can recover an estimate of the mean WTP (assuming the mean of the proposed distribution is finite).[8] The binary vote model that is implied by the population WTP distribution together with a finite set of tax amounts can then be used to describe the implied distribution of votes at each tax amount. Maximum likelihood estimation is feasible for standard distributions. The primary drawback with this type of parametric approach arises because the estimate of the mean WTP can be sensitive to the chosen distribution. For these parametric models it is usually not possible to determine whether the estimate for the mean WTP will tend to underestimate or overestimate the true population mean.[9]

Parametric models are useful for describing, in a simple way, the relationships between the vote (represented here as *Y)* and observable covariates, including the tax amount, income, and attitude variables – even though they might not be reliable for estimating mean WTP. Such binary response models – with the leading contenders in empirical work being the probit, logit, and linear probability models – do not have to be correctly specified in order to provide useful information. In fact, as discussed in Angrist and Pischke (2009) and Wooldridge (2010, Chapter 15), and as empirical economic research over many years has verified, the linear probability model usually provides reliable estimates of directions and magnitudes of the effects of independent variables, on average. See also the discussion in Appendix 1.4. Therefore, the linear probability model is the primary basis for evaluating the role of adding explanatory variables to the description of the determinants of respondents' votes for the prevention program. As a further robustness check, probit models were also estimated. These results are reported in Appendix 6.1.

10.4.2 OLS ANALYSIS OF THE EFFECTS OF TAX AMOUNT, INCOME, AND ATTITUDES

This analysis is undertaken for the Version A and Version B injury descriptions separately. These models focus on four types of variables: (a) the randomly assigned tax amount; (b) the family income and the dummy variables that account for the way respondents answered the family income question and the year when the income was reported[10]; (c) attitude measures elicited before the injuries and prevention plan were explained to each respondent, as proxy measures for respondents' tastes[11]; and (d) fixed effects for the residential zip code of each respondent. The motivation for including zip

[7] The specific transformation of each independent variable will affect the interpretation of a unit change.

[3] See Technical Memo TM-3 for the definition of the willingness to pay for the prevention program and discussion of the assumptions underlying the derivation of the willingness to pay distribution. See Appendix 1.4 for a more detailed discussion of using parametric models to estimate mean WTP.

[9] In statistical parlance, the direction of the bias cannot be determined.

[10] All of the analyses reported in this technical memo use the imputed values for the family income. Appendix 1.1 describes the income imputation procedure used.

[11] Many conventional models of consumer behavior would argue it is important to control for individual taste variables as indicators of differences in preferences in order to interpret the estimated income effects.

code effects is that, while the tax amount is randomized and therefore independent of both observed and unobserved taste variables, income is clearly not randomized (and it is self reported). The inclusion of zip code effects recognizes that average incomes vary by zip code as do unobserved taste variables that might be expected to also affect the vote. The zip code effects allow the analysis to estimate the effect of income on the vote while controlling for other unobserved taste variables. By including the observed taste variables and the zip code effects the analysis accounts indirectly for differences in individual tastes that are correlated with income.[12]

Tables 6 and Table 7 present the OLS estimates of five models for respondents' votes for each injury scenario. Each column in the table reports a different specification for the hypothesized determinants of the votes for the prevention program. The first element in each row in the table reports the OLS estimate for the coefficient associated with each independent variable. The names of these variables are given in the far left column of the table. The statistics below each estimated coefficient (in parentheses) are p-values for the two-sided test that the unknown the coefficient of each variable is zero. The number of observations in the sample used to estimate the model is in the last row of the table, and the R-squared measure of the "goodness of fit" is in the row above it. Each table also reports whether or not the specification includes the zip code fixed effects. Finally, the ratio of the coefficient for the logarithm of income to the absolute magnitude of the coefficient for the logarithm of the tax amount is reported in the third row from the bottom of the table along with the p-value for the null hypothesis testing whether the elasticity was zero. This ratio is labeled the 'income elasticity.' See Appendix 1.7 for discussion of how the elasticity of WTP with respect to income can be obtained from a parametric binary response model for vote.

[12] The Tiebout [1956] hypothesis has received wide support in urban and environmental economic applications (see Kuminoff et. al.[2013] for a review) This framework hypothesizes that individuals select places to live considering the associated labor markets and jobs as well as the local public services and amenities.

TABLE 6. OLS ESTIMATES FOR LINEAR PROBABILITY MODEL FOR VOTES: VERSION A

OLS Estimator (Vote)	(1)	(2)	(3)	(4)	(5)
Log of tax amount	**-0.085**	**-0.085**	**-0.085**	**-0.089**	**-0.087**
	0.000	0.000	0.000	0.000	0.000
Log family income		0.016	**0.039**	**0.029**	**0.051**
		0.180	0.030	0.011	0.005
Year 2013 income elicited		0.015	0.033	0.018	0.023
		0.439	0.320	0.339	0.462
$0 income response		0.183	**0.449**	0.260	**0.552**
		0.217	0.041	0.089	0.016
Income category question answered				-0.020	0.026
				0.459	0.468
Income bounding question answered				**-0.143**	-0.075
				0.002	0.256
Respondent refused income question				**-0.167**	-0.115
				0.000	0.068
Importance of protecting coast				**0.067**	**0.081**
				0.000	0.000
Importance of reducing federal taxes				**-0.043**	**-0.037**
				0.000	0.004
Spending on cleaning up pollution				**0.080**	**0.061**
				0.000	0.000
Constant	**0.767**	**0.588**	0.386	0.066	-0.212
	0.000	0.000	0.072	0.664	0.359
Zip Fixed Effects	No	No	Yes	No	Yes
Income elasticity		0.185	**0.456**	**0.324**	**0.584**
P-value		0.193	0.048	0.016	0.011
R-squared	0.043	0.045	0.323	0.125	0.372
Observations	1,833	1,833	1,833	1,833	1,833

Note: P-values appear underneath coefficients. Coefficients in **bold** significant at the 0.05 level.

The first column in Table 6 reports the simplest specification with the determinant of a respondent's vote limited to the log of the tax amount he received. These results are consistent with the pair-wise tests of the cost condition. Namely, increases in the tax amount have a negative effect on the probability of voting for the prevention program. The estimated magnitude of the coefficient implies that a 10 percent increase in the tax amount lowers the vote percentage by slightly less than one percentage point (specifically a decrease of .85 percentage points). Comparing the magnitude of the estimated coefficient for the log of the tax amount across alternative model specifications, it is clear that the implied effect of this variable on the probability of voting for the program is stable. This is as expected because the tax amount is randomly assigned, and therefore independent of observed and unobserved factors affecting the vote.

The models presented in the next four columns, labeled (2) thru (5), progressively include additional variables. Column (2) includes the log of income along with two dummy variables associated with income. Thirty-seven respondents (18 for the Version A injury description and 19 for Version B) reported zero income. To allow for the logarithmic transformation – which allows simple calculation of an elasticity measure with respect to income – zero income responses were assigned a value of one. Thus, the value for the log of income for these observations is zero. A dummy variable labeled "0$ income response" is included to adjust for this assignment. From the estimates it can be seen that individuals reporting zero income, who comprise about one percent of the total sample, vote in favor of the program at high rates. Omitting these individuals does not change the substantive findings for the coefficients on the other variables. The second dummy variable is assigned unity for the family income reported for 2013 and zero for 2012 (1,163 respondents were interviewed in 2014 so the family income was for 2013 with Version A and 1,124 for Version B). This dummy variable is generally statistically insignificant.

Column (3) uses the same model specification and includes zip code fixed effects. The coefficient for the log of family income with this model specification is significantly different from zero. Column (4) includes the dummy variables for the ways respondents could have answered the income question, as described in Technical Memo TM-13.[13] A dummy variable is also included for cases where the respondent did not answer the income question. Finally, three attitude variables, asked before the question reporting each respondent's vote decision, are included.[14] The omitted category for the income response variable is the open-ended response. The coefficients for the income response dummy variables reflect the adjustment associated with the effects of imputation compared to a direct report of family income. Zip code fixed effects are not included in this specification. The log of income is statistically significant. All of the attitude variables are also significant determinants of the votes with signs consistent with interpreting them as reflecting respondents' tastes. For example, positive attitudes toward protecting coastal areas and cleaning up pollution are associated with positive conditional effects on the probability of voting for the prevention program.

The last model in column (5) repeats this specification and includes the zip code fixed effects. The general conclusions are unchanged. The model fit improves as expected. The tax amount and family income variables are statistically significant determinants of the respondents' probability of voting for the program. The sign of the tax amount is consistent with the cost condition and the sign of log income with conventional economic expectations when income is interpreted as a proxy for ability to pay.

Table 7 reports the estimates for the same five model specifications using the sample that was randomly assigned to Version B's injury description. The organization of the table is

[13] Appendix 6.1 reports alternative model specifications, including different treatments of family income and interactions of the dummy variables for income responses with the log of income.

[14] When the questions underlying these attitude variables were not answered, the missing values were imputed using the procedures described in Appendix 1.1. Estimates of these models with samples restricted to respondents who answered the attitude questions are in Appendix 6.1.

the same as Table 6. The estimated effect for the log of the tax amount is negative and significantly different from zero. The magnitudes are similar to those found in Table 6. The coefficient for the log of income is positive and is significantly different from zero, with small p-values, regardless of the model specification. The effects of the attitude variables are consistent with the results for the models estimated using the sample receiving the Version A injury description.

TABLE 7. OLS ESTIMATES FOR LINEAR PROBABILITY MODEL FOR VOTES: VERSION B

OLS Estimator (Vote)	(1)	(2)	(3)	(4)	(5)
Log of tax amount	**-0.088** 0.000	**-0.088** 0.000	**-0.086** 0.000	**-0.095** 0.000	**-0.092** 0.000
Log family income		**0.042** 0.000	**0.046** 0.002	**0.046** 0.000	**0.045** 0.002
Year 2013 income elicited		**-0.050** 0.046	-0.048 0.203	-0.026 0.274	-0.035 0.328
$0 income response		**0.617** 0.000	**0.578** 0.006	**0.575** 0.000	**0.526** 0.017
Income category question answered				**-0.068** 0.009	-0.060 0.087
Income bounding question answered				**-0.179** 0.002	**-0.201** 0.027
Respondent refused income question				**-0.260** 0.000	**-0.223** 0.001
Importance of protecting coast				**0.064** 0.000	**0.054** 0.003
Importance of reducing federal taxes				**-0.060** 0.000	**-0.056** 0.000
Spending on cleaning up pollution				**0.087** 0.000	**0.088** 0.000
Constant	**0.829** 0.000	**0.414** 0.000	**0.419** 0.022	0.052 0.710	0.087 0.687
Zip Fixed Effects	No	No	Yes	No	Yes
Income elasticity P-value		**0.472** 0.000	**0.531** 0.005	**0.487** 0.000	**0.488** 0.004
R-squared	0.045	0.055	0.301	0.158	0.364
Observations	1,823	1,823	1,823	1,823	1,823

Note: P-values appear underneath coefficients. Coefficients in **bold** significant at the 0.05 level.

The last estimated parameter in each column of Table 6 and Table 7 (reported above the row for R^2 for each model) is labeled as the "income elasticity." The assumptions

required to interpret this elasticity as an income elasticity of WTP are discussed in Appendix 1.7. This elasticity is easily computed as

$$-\frac{\gamma}{\alpha}, \tag{1}$$

where α is the coefficient on $\log(Tax)$ (where Tax is the tax, amount) and γ is the coefficient on $\log(Income)$ in the LPM equation

$$P(Y = 1|X, B) = \alpha \log(Tax) + \gamma \log(Income) + X\beta. \tag{2}$$

As noted, Appendix 1.7 describes how the elasticity is obtained and Hanemann [1999,pages 65-66] contains a more detailed discussion.

As can be seen in Table 6, for Version A the estimated income elasticity ranges from .185 to .584, and is statistically different from zero at the five percent level in three of the four cases. Such variation in the elasticity estimate would be expected as explanatory variables are added to the equation: income is not randomly assigned and is likely correlated with unobserved tastes (including those captured by the zip code effects). In addition income is self- reported. The largest elasticity estimate is found for the model reported in column (5) when the full set of observed taste variables and zip code fixed effects are included.

For Version B, the estimated income elasticity is more stable, ranging from .472 to .531, and it is always statistically significant with p-value well below .01. Overall, once the one percent of respondents reporting zero income are accounted for in the model specification and proxy measures for taste variables are included, the estimated elasticity is approximately .5 for Versions A and B.

10.4.3 BACKGROUND FOR NONPARAMETRIC ESTIMATORS OF THE MEAN WTP

The linear probability model describes how the randomly assigned tax amount and variables describing each respondent, such as his family income or attitudes toward environmental programs, influence his vote for the prevention program. It is also possible to model the mean willingness to pay implied by these votes directly. An estimator based on this approach is the basis for the second set of multivariate models presented here.

Before describing the results with this estimator some context is desirable. Lewbel (2000) and Watanabe (2010) have demonstrated that the mean WTP can, under certain assumptions, be consistently estimated without parametric assumptions. One useful result in the Lewbel/Watanabe (LW for short) approach is a simple relationship between mean WTP and the statistical concept known as the "survival function" for the WTP distribution.[15] With random assignment of the tax amount, it is possible to estimate the mean WTP from a sample with the tax amounts and the resulting binary votes. This result for the LW approach has attractive features and drawbacks. The LW approach requires a continuous design for the tax amounts as well as a maximum tax amount that exceeds the maximum WTP in the population. Under these conditions, the LW approach provides an

[15] See Technical Memo TM-13 for definitions and a description of the key probabilistic results. The function describing the probability that a continuous random variable is less than a specific value for a realization of that variable is labeled the cumulative probability function. The function describing the probability that variable is at least that realized value is labeled the survival function.

unbiased and consistent estimate of the mean WTP. However, the requirements to establish these results are problematic for practical purposes. A continuous design for the tax amounts proposed implies that each individual is very likely to receive a different tax amount. Under these conditions each value of the survival function is estimated imprecisely.[16] Imprecise estimates of this function make testing of the cost condition for consistent economic choices in the population problematic. The second requirement, that the maximum tax amount is above all WTPs in the population, is infeasible. There is no a priori basis for knowing the maximum WTP in the population.

It is possible to relax both requirements in implementing the LW approach. Both a discrete design for the tax amounts and a maximum value below the maximum WTP in the population can be used in implementing the method. These modifications produce a *conservative* estimator of the mean WTP.[17]

When the estimated survival function satisfies the monotonicity requirement, the LW estimate is identical to the ABERS estimate [Ayer et al. (1955)] that has been used in a number of applications of contingent valuation. There are no monotonicity violations in the samples for Version A or Version B injury descriptions. Thus the two approaches yield identical results for this analysis. Lewbel (2000) and Watanabe (2010), along with Lewbel, Linton, and McFadden (2011), have also demonstrated how to incorporate covariates. This extension corresponds to the second set of estimates discussed below.

10.4.4. THE LEWBEL-WATANABE ESTIMATOR WITH COVARIATES

The LW approach can consider cases where only the (conditional) mean function in the WTP distribution needs to be specified. Both Lewbel (2000) and Watanabe (2010) proposed linear models for E($WTP|X$), where X corresponds to a vector of variables hypothesized to influence the conditional mean WTP. Linear models have the shortcoming that they can predict negative WTP for at least some subsets of the population described by X. Therefore the multivariate models used for this analysis adopt an exponential specification as in equation (2).

$$E(WTP|X) = \exp(X\beta), \qquad (3)$$

where β is a vector of parameters multiplying each variable in the row vector X, including an intercept. The exponential formulation implies that the coefficient β_j for the independent variable, X_j, is the elasticity of the mean WTP with respect to variable associated with X_j, when X_j is in logarithmic form.

[16] Estimates for this function must make functional form assumptions or encounter imprecise and irregular shapes for the resulting empirical functions.

[17] The details are described in Appendix 1.4. With a discrete design for the tax amounts, the LW estimator is easily obtained from a linear regression of the binary vote outcome on binary variables indicating the different tax amounts. The estimate of the lower bound mean is then a linear combination of the estimated coefficients in the linear regression (which are simply the proportion of yes votes at each tax amount). Obtaining a lower bound estimator via linear regression is convenient for obtaining reliable standard errors, as standard methods for accounting for survey design can be applied directly.

Using a discrete distribution for the tax amounts, with a tax amount less than the maximum WTP in the population, implies that the LW theoretical results for the properties of the estimates for the parameters β do not apply. As a result, to evaluate the properties of the estimator under these conditions a set of simulation experiments was conducted using frameworks and distributional assumptions that were hypothesized to match those of the final survey. The results from those experiments indicated that the coefficients for continuous variables, such as income, had a systematic downward bias when the design for the tax amounts was discrete.[18]

As described in Appendix 1.4, the LW estimates are obtained by weighting each vote by the inverse of the proportion of observations at the particular tax amount. The estimator adopts a Poisson regression approach.[19]

Tables 8 and 9 provide the estimates for the four model specifications coinciding with the binary response linear probability models – (2) thru (4) in Table 6 and 7 (aside from including the tax amount variable) for the samples associated with the Version A and Version B injuries respectively. The model in column (1) of Table 8 includes only the income, in logarithmic form, and the dummy variables for zero income and for the year when family income was requested. The income elasticity is not statistically significant and the estimated magnitude is small. The magnitude increases to 0.121 for the most complete specification for the model in column (4). This specification includes the income variables, dummy variables for the income response terms, and the taste variables along with zip code fixed effects. The income elasticity is statistically significant for this specification with p-value = 0.035. The magnitude of the elasticity is stable across the specifications reported in columns (2) thru (4) but the elasticity is only significant (using two sided criteria) for the model in column (4). The attitude variables are significant determinant of WTP and their effects are consistent with a priori expectations and what was found with the linear probability model.

[18] See Appendix 1.3 for the simulation results.

[19] The approach relies on a well-known result from the generalized linear models (GLM) and quasi-maximum-likelihood estimation (QMLE) literature on the robustness of this estimator to arbitrary distributional misspecification. In addition, as discussed in Appendix 1.4, the choice of an exponential mean function and the Poisson QMLE yields numerically identical estimates of the lower bound mean WTP reported below in Table 9 when the independent variables serving as covariates in each model specification are averaged out. This assumes that the same set of observations is used for the simple estimate and the estimate obtained with covariates.

TABLE 8. LEWBEL-WATANABE ESTIMATOR WITH COVARIATES FOR VERSION A

Poisson Estimator (Vote/Tax Density)	(1)	(2)	(3)	(4)
Log family income	0.041 0.382	0.103 0.075	0.082 0.073	**0.121** 0.035
Year 2013 income elicited	0.071 0.385	0.105 0.333	0.089 0.271	0.063 0.562
$0 income response	-0.071 0.912	0.380 0.614	0.217 0.736	0.594 0.448
Income category question answered			0.048 0.635	0.177 0.145
Income bounding question answered			-0.483 0.057	-0.369 0.270
Respondent refused income question			**-0.807** 0.000	**-0.808** 0.001
Importance of protecting coast			**0.295** 0.000	**0.343** 0.000
Importance of reducing federal taxes			**-0.144** 0.000	**-0.145** 0.001
Spending on cleaning up pollution			**0.272** 0.000	**0.180** 0.000
Constant	**4.400** 0.000	**4.018** 0.005	**2.169** 0.001	1.955 0.179
Zip Fixed Effects	No	Yes	No	Yes
Observations	1,833	1,833	1,833	1,833

Note: P-values appear underneath coefficients. Coefficients in **bold** significant at the 0.05 level.

The significant effect of family income is consistent for all specifications using the Version B injury description. It is estimated to be between .138 and .150. All of the estimates are statistically significant with p-values below 0.01. When the full model specification is considered –as given in column (4) in each table - there is only a modest difference in the income elasticity between the A and B injury descriptions.

TABLE 9. LEWBEL-WATANABE ESTIMATOR WITH COVARIATES FOR VERSION B

Poisson Estimator (Vote/Tax Density)	(1)	(2)	(3)	(4)
Log family income	**0.138** 0.000	**0.145** 0.010	**0.150** 0.000	**0.150** 0.011
Year 2013 income elicited	-0.147 0.074	-0.107 0.315	-0.065 0.425	-0.031 0.772
$0 income response	**1.983** 0.000	**1.910** 0.006	**1.880** 0.000	**1.922** 0.012
Income category question answered			-0.153 0.076	-0.134 0.218
Income bounding question answered			-0.433 0.127	-0.627 0.059
Respondent refused income question			**-0.857** 0.006	**-0.857** 0.014
Importance of protecting coast			**0.215** 0.000	**0.218** 0.001
Importance of reducing federal taxes			**-0.164** 0.000	**-0.170** 0.000
Spending on cleaning up pollution			**0.276** 0.000	**0.286** 0.000
Constant	**3.634** 0.000	3.777 0.078	**2.118** 0.000	2.170 0.293
Zip Fixed Effects	No	Yes	No	Yes
Observations	1,823	1,823	1,823	1,823

Note: P-values appear underneath coefficients. Coefficients in **bold** significant at the 0.05 level.

10.5 SCOPE TESTS

As noted in section 10.3, Figure 1 indicates that the quantity condition is met in our data. In other words, the fact that at each of the tax amounts the vote percentage for, Version B, is above the corresponding vote percentage for Version A, suggests that it is not possible to reject the quantity condition. Thus, the survey data are consistent with the quantity condition, as they were with the cost condition.

As in testing the cost condition, it is possible to use more stringent tests of the quantity condition. These tests are often described as the scope tests. In each case, the null hypothesis is that the quantity condition does not hold and the alternative is that the strong form of the quantity condition holds. For example, as discussed in Appendix 1.4, one can take the null hypothesis to be $E(WTP_A) \geq E(WTP_B)$ – the mean WTP for the program under Version B is no greater than that for Version A – in which case the alternative is $E(WTP_B) > E(WTP_A)$. If the estimate of $E(WTP_B)$ is less than $E(WTP_A)$ then the null hypothesis is never rejected. The null is rejected in favor of the one-sided alternative only if the estimate of $E(WTP_B)$ is significantly above $E(WTP_A)$, where significance is determined by a suitable p-value. Exactly the same approach can be used

for comparing the percentages of yes votes across each of the two scenarios (Versions A and B) and across the five tax amounts.

Table 10 summarizes the results for two scope tests using OLS estimates for six different specifications of the linear probability model. These analyses use a sample that pools the responses for the questionnaires associated with the Version A and Version B injury descriptions. Column (1) of the table provides the simplest and most direct test of scope. Since the tax amount and version associated with the injury description were randomly assigned, a simple regression equation with an intercept and a dummy variable for the version (with those respondents receiving the questionnaire corresponding to the Version B injuries = 1 and those with the Version A = 0) provides a direct estimate of the difference in the probabilities of voting for the program in the Version B injury description and the Version A injury description. The t statistic provides us with the scope test.

The estimated effect for the version dummy variable is significant with a p-value for a one-sided test of 0.002. The estimated coefficient implies approximately a five percentage point increase in the probability of voting for the prevention program with the Version B injuries compared with the Version A injuries. Thus, the survey provides support for the strong form of a scope test.

Because of random assignment of both the version and tax amounts, adding indicators for the different tax amounts assigned to the respondents should not appreciably change the scope test. It is possible to develop more precision in the estimated scope effect because the tax amounts help to predict vote. The second column of Table 10 indicates that the estimated scope effect, .047, is the same to three decimal places as that in column (1). The p-value is smaller (.001), reflecting the reduction in error variance in the equation with inclusion of the tax amount dummies. It should be noted that the regression in column (2) does not include an intercept because a full set of tax amount dummies was included.

The second scope test compares the difference in votes for the prevention program for the two injury descriptions at each tax amount separately. It is a more demanding test because the sample sizes are reduced when the respondents assigned to each tax amount but different versions are considered distinct sub-samples. These hypotheses can be tested using interaction dummy variables. Each variable is the product of the dummy variable for each tax amount with the dummy variable defining version. The third column in Table 10 reports these estimates. All of the interaction effects are positive, reflecting the results in Figure 1: the survival function for the Version B injuries is above that for the Version A injuries. Three of the five interaction coefficients are significantly different from zero with p-values (rounded to two decimal places) no greater than 0.10. Two of the p-values are less than 0.05. When zip code fixed effects are included it is possible to reduce the amount of unexplained variation in the vote. The estimates are reported in column (4). Now four of the five interaction coefficients are significantly different from zero at p-values of 0.10 or smaller (using one sided alternatives).

Rather than considering the difference in vote percentages at each tax amount in columns (3) and (4), it is possible to use a joint test of whether all coefficients on the interaction

terms are zero. The results are given in the row labeled "Interaction Terms Jointly = 0." In column (3) the p-value for this test is .07. When zip code fixed effects are added the p-value falls to .04. Thus, the joint test is consistent with the one-sided t tests. The joint test is inherently two sided, and so it is a conservative test of scope.

The tests in columns (1) and (2) assume nothing in particular about the shape of the survival functions for the Version A and Version B injury descriptions. Instead, they are simply testing scope based on the overall vote probabilities. Nevertheless, it is of some interest to test the null hypothesis that the survival function for Version B is a constant shift from the survival function in Version A. Figure 1 indicates that a constant shift provides a reasonable description. A formal statistical test would consider whether the interaction terms in columns (3) and (4) are all the same. In other words, the model estimated in column (2) is the restricted version of the model estimated in column (3) (since neither of these equations contains other explanatory variables). The p-value for the test that the survival functions differ by a constant amount is given in the row labeled "Interaction Terms Equal." In column (3) the p-value (against a two-sided alternative) is .92 and when zip code effects are added p-value = .88. This is essentially no evidence to contradict the conclusion that the survey respondents' votes on average are consistent with the quantity condition. The null hypothesis that the survival function for Version B is a constant shift up – almost five percentage points – from the survival function for Version A.

Columns (5) and (6) consider whether the conclusion of the basic scope test changes change if the tax amount, family income, and taste variables in the analysis of the samples for each version of the injury description are included in the model using conventional functional forms. There is no change in the conclusions. Both the test results and the magnitude of the effect of version remain approximately the same with a five percentage point increase in the probability of voting for the prevention program with the Version B injury description. Overall, Table 10 demonstrates the robustness of the scope test results.

TABLE 10. OLS ESTIMATES FOR LINEAR PROBABILITY MODEL FOR SCOPE

OLS Estimator (Vote)	(1)	(2)	(3)	(4)	(5)	(6)
Version B*	**0.047**	**0.047**			**0.050**	**0.062**
	0.002	0.001			0.000	0.000
Tax Amount 15		**0.526**	**0.522**	**0.604**		
		0.000	0.000	0.000		
Tax Amount 65		**0.438**	**0.435**	**0.514**		
		0.000	0.000	0.000		
Tax Amount 135		**0.345**	**0.356**	**0.433**		
		0.000	0.000	0.000		
Tax Amount 265		**0.291**	**0.283**	**0.358**		
		0.000	0.000	0.000		
Tax Amount 435		**0.238**	**0.242**	**0.313**		
		0.000	0.000	0.000		
Version B 15*			**0.055**	**0.065**		
			0.040	0.034		
Version B 65*			0.053	0.068		
			0.082	0.061		
Version B 135*			0.024	0.023		
			0.246	0.280		
Version B 265*			**0.063**	**0.073**		
			0.029	0.029		
Version B 435*			0.039	**0.062**		
			0.100	0.037		
Log of tax amount					**-0.090**	**-0.091**
					0.000	0.000
Log family income					**0.041**	**0.056**
					0.000	0.000
Year 2013 income elicited					-0.008	0.001
					0.591	0.976
$0 income response					**0.477**	**0.611**
					0.000	0.000
Importance of protecting coast					**0.066**	**0.066**
					0.000	0.000
Importance of reducing federal taxes					**-0.052**	**-0.048**
					0.000	0.000
Spending on cleaning up pollution					**0.086**	**0.075**
					0.000	0.000
Constant	**0.368**				-0.041	-0.177
	0.000				0.704	0.174
Zip Fixed Effects	No	No	No	Yes	No	Yes
Interaction Terms Jointly = 0						
Wald Test F-statistic			2.066	**2.375**		
P-value			0.070	0.040		
Interaction Terms Equal						
Wald Test F-statistic			0.231	0.293		
P-value			0.921	0.883		
R-squared	0.002	0.420	0.420	0.509	0.130	0.254
Observations	3,656	3,656	3,656	3,656	3,656	3,656

Note: P-values appear underneath coefficients. Coefficients in **bold** significant at the 0.05 level.
* Starred coefficients' p-values are for one-sided test that value is greater than zero.

As noted in Technical Memo TM-3 and Appendix 1.4, a common approach for testing the scope effect considers the difference in the lower bound willingness to pay estimate for the prevention program using the samples corresponding to the different injury descriptions separately. It is based on the economic logic developed in Technical Memo TM-3. As discussed in Appendix 1.4, using estimates of lower bound means raises difficulties in testing for scope. The problem arises because both lower bound mean estimators have a downward bias, and without information about the WTP distributions in the population it is not possible to assess whether the difference in lower bound means is greater than or less than the difference in population mean WTP. As discussed further in Appendix 1.4, and as illustrated by an example, there are reasons to believe that with low tax values (as in the current study), the lower bound mean for the Version B scenario has a larger downward bias than that for the Version A scenario. In this case, the estimated scope effect is conservative, and uncovering a scope effect would be more difficult. Because both estimators have a downward bias – whose magnitude cannot be determined in advance.

In any case, this survey used the same five tax amounts for both injury descriptions. To obtain a scope test based on mean WTP that does not impose parametric assumptions, there is no alternative option but the use of the lower bound estimators.

Table 11 reports the lower bound mean for each sample and a test for equality of these means. As the p-values suggest, the test rejects the null hypothesis of equality of means with a p-value below 0.01.

TABLE 11.

Injury Description	Ayers et al Lower Bound Mean WTP	Standard Error	t	One-sided P-value	95% confidence interval	
Version A	132.36	5.52	23.99	0.000	121.49	143.23
Version B	152.25	6.04	25.23	0.000	140.37	164.14
B-A	19.89	7.55	2.63	0.004	5.02	34.77

The last component of the analysis of scope applies the LW estimator, as described above, with the pooled sample including Version A and Version B injury descriptions. Table 12 reports four specifications. Each of the equations includes the version dummy variable, which can be given a percentage interpretation (when multiplied by 100). Model (1) includes the log of income along with dummy variable for a zero income response and for the year when income was elicited. Model (2) includes, in addition to the income variables, attitude variables as proxy measures of respondents' tastes. The third and fourth models include the income response dummy variables in (3) and the zip code fixed effects are added to that specification in (4). Across the four specifications, the magnitude of the scope effect ranges from about 13.6 percent to 16.3 percent, and all estimates are statistically different from zero with two-sided p-values less than .01. The largest effect is obtained when income, taste variables, and zip code effects are controlled for.

It is possible to compare the estimated percentage effects in Table 12 with the percentage change in lower bound means for in Table 11. The Version A estimate is $132.36 and the Version B estimate is $152.25, so the Version B estimate is about 15 percent higher than the Version A estimate. This difference accords well with the results in Table 12, providing further evidence for the robustness of the scope effect.

TABLE 12. LEWBEL-WATANABE ESTIMATOR COVARIATES POOLED WITH VERSION EFFECT

Poisson Estimator (Vote/Tax Density)	(1)	(2)	(3)	(4)
Version B	**0.140** 0.009	**0.148** 0.004	**0.136** 0.009	**0.163** 0.004
Log family income	**0.094** 0.002	**0.121** 0.000	**0.120** 0.000	**0.154** 0.000
Year 2013 income elicited	-0.047 0.431	-0.005 0.934	0.007 0.906	0.048 0.493
$0 income response	**1.161** 0.003	**1.307** 0.001	**1.244** 0.002	**1.520** 0.001
Importance of protecting coast		**0.252** 0.000	**0.251** 0.000	**0.267** 0.000
Importance of reducing federal taxes		**-0.158** 0.000	**-0.157** 0.000	**-0.151** 0.000
Spending on cleaning up pollution		**0.278** 0.000	**0.277** 0.000	**0.252** 0.000
Income category question answered			-0.056 0.381	-0.035 0.646
Income bounding question answered			**-0.466** 0.018	**-0.434** 0.047
Respondent refused income question			**-0.833** 0.000	**-0.914** 0.000
Constant	**3.908** 0.000	**1.993** 0.000	**2.065** 0.000	1.758 0.176
Zip Fixed Effects	No	No	No	Yes
Observations	3,656	3,656	3,656	3,656

Note: P-values appear underneath coefficients. Coefficients in **bold** significant at the 0.05 level.

REFERENCES

Angrist, Joshua D. and Jorn-Steffen Pischke, 2009 Mostly Harmless Econometrics: An Empiricist's Companion (Princeton,N.J.: Princeton University Press)

Ayer, Miriam, H. D. Brunk, G.M. Ewing, W.T. Reid, and Edward Silverman. 1955. "An Empirical Distribution Function for Sampling with Incomplete Information." *The Annals of Mathematical Statistics* 26(4): 641–647.

Hanemann, W. Michael,1999. "Neo-classical Economic Theory and Contingent Valuation" in Ian J. Bateman and Kenneth G. Willis, editors, Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries (New York: Oxford University Press) pp.42-96.

Hanemann, W. Michael and Barbara Kanninen,1999. "The Statistical Analysis of Discrete –Response CV Data" in Ian J. Bateman and Kenneth G. Willis, editors, Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries (New York: Oxford University Press) pp.302-441.

Kuminoff, Nicolai V., V. Kerry Smith and Christopher Timmins, 2013, "The New Economics of Equilibrium Sorting and its Transformational Role for Policy Evaluation." *Journal of Economic Literature*,51(4):1-57.

Lewbel, Arthur, 2000, "Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables" Journal of Econometrics, 97 (): 145-177.

Lewbel, Arthur, Daniel McFadden, and Oliver Linton, 2011, "Estimating features of a distribution from binomial data," Journal of Econometrics 162(): 170-188.

Tiebout, Charles M. (1956). "A Pure Theory of Local Expenditures." *Journal of Political Economy*, 64(5) October: 416-424.

Turnbull, B.W. 1976, "The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data," Journal of the Royal Statistical Society 38( ): 290-295.

Watanabe, Masahide, 2010, "Nonparametric Estimation of Mean Willingness to Pay from Discrete Response Valuation Data," American Journal of Agricultural Economics 92 (4): 114-1135.

Wooldridge, Jeffrey M., 2010, Econometric Analysis of Cross Section and Panel Data, second edition, (Cambridge,Mass: The MIT Press).