Attachment 15: Description of Statistical Survey Design

The following represents an anticipated experimental design for survey implementation, along with the associated number of completed surveys that will be required. Part B of this supporting statement provides detail on the sampling design. The proposed design and sampling plan is based on standard design and sampling theory for choice experiments and population surveys, as outlined by Louviere et al. (2000), Kuhfeld (2009) and Dillman (2000). EPA notes that the anticipated experimental design described here is preliminary and it may be subject to refinements during design evaluations to account for issues such as dominant or dominated pairs, ecological feasibility, and to remove attribute combinations which do not provide information for estimation.

The purpose of the Chesapeake Bay survey is to calculate average per household parameters (e.g., willingness to pay and choice probabilities) within a given survey population. Additional analysis that differentiates per-household parameters may be conducted within groups of households which use or do not use the Chesapeake Bay.

Experimental design for the choice experiments

Based on focus groups and pretests, and guided by realistic ranges of attribute outcomes, the anticipated experimental design includes a fixed status quo or "no policy" option (Option A), and two multi-attribute choice options or *alternatives*, Options B and C. These alternatives, Option B and Option C, are characterized by three potential levels for environmental attributes and six different levels of annual household cost. The study design will employ two split samples. The first consists of three different baselines, where Option A or the status quo choice represents either a constant, declining, or improving baseline in the environmental attributes, relative to current levels. The second experimental treatment involves two different reference years for predictions of the attribute levels (2025 or 2040).

Each of these split-sample experiments will be produced in the three geographic divisions:

1. Bay States: Maryland, Virginia, District of Columbia
2. Watershed States: Delaware, New York, Pennsylvania, and West Virginia
3. Other East Coast States: Vermont, New Hampshire, New Jersey, Massachusetts, Connecticut, Rhode Island, Maine, North Carolina, South Carolina, Georgia, and Florida

Within these geographic divisions, the split-sample experiment cells will collect more detailed information in the Bay states, as stipulated in Table A15-1. Based on the three geographic strata, three baseline versions, and two reference years for environmental conditions, the study will consist of the following 18 cell design:

Table A15-1. Split-sample design cells.

	Geographic division	Baseline factor	Reference year for environmental conditions
Cell 1	Bay States	Declining	2025
Cell 2	Bay States	Constant	2025
Cell 3	Bay States	Improving	2025
Cell 4	Watershed	Declining	2025
Cell 5	Watershed	Constant	2025
Cell 6	Watershed	Improving	2025
Cell 7	Other East Coast	Declining	2025
Cell 8	Other East Coast	Constant	2025
Cell 9	Other East Coast	Improving	2025
Cell 10	Bay States	Declining	2040
Cell 11	Bay States	Constant	2040
Cell 12	Bay States	Improving	2040
Cell 13	Watershed	Declining	2040
Cell 14	Watershed	Constant	2040
Cell 15	Watershed	Improving	2040
Cell 16	Other East Coast	Declining	2040
Cell 17	Other East Coast	Constant	2040
Cell 18	Other East Coast	Improving	2040

Options B and C, the alternatives to the status quo baseline (option A), are characterized by different levels for the following six attributes, including cost. The number of levels corresponding to each of these attributes is depicted below:

1. Change in water clarity in Options B and C (x_{1B}; x_{1C}) – 3 possible levels

2. Change in adult blue crab abundance in Options B and C (x_{2B}; x_{2C}) – 3 possible levels
3. Change in adult striped bass abundance in Options B and C (x_{3B}; x_{3C}) – 3 possible levels
4. Change in oyster abundance in Options B and C (x_{4B}; x_{4C}) – 3 possible levels
5. Change in lake condition in Options B and C (x_{5B}; x_{5C}) – 3 possible levels
6. Cost in Options B and C (x_{6B}; x_{6C}) - 6 possible levels

This implies an experimental design characterized by three levels for each of the five attributes and six levels for costs [$3^5 \times 6$] for each alternative option, or [$3^{10} \times 6^2$] for Options B and C together.

To construct a preliminary main effects design with 72 profiles that is sufficiently flexible to estimate alternative specific main effects and response patterns (i.e., a non-generic design), we begin with a 3^5x6 orthogonal fractional factorial design with 144 profiles. We then combined the elements of this design into pairs that would reflect trade-offs at the margin (i.e., improvements in the attributes that are attained at the cost of decrease in other environmental attributes and/or increase of the overall cost of the program). Finally, these pairs were blocked[1] in such a way that variability of the environmental and cost attributes within a block would be maximized (and hence the main effects would not be confounded with the block effects). The result is a design with 72 profiles, with attributes labeled following the above notation, and levels indicated by integers 1...N, where N for each attribute is the number of levels identified above.

Following common practice in the environmental economics literature, we anticipate three choice questions per survey. This allows the 72 profiles to be included (orthogonally blocked) in 24 unique survey booklets, as illustrated in Table A15-2. The attribute levels applied within surveys are summarized in Table A15-3. Monte Carlo evidence suggests that 6 to 12 completed responses are required for each profile in order to achieve large sample statistical properties for choice experiments (Louviere et al. 2000, p. 104, citing Bunch and Batsell 1989). Following this guidance, the above design will require 24×12 = 288 completed surveys, or 12 completed surveys for each unique survey booklet,. This will provide a total of 864 profile responses per cell.

Table A15-2. The set of 72 design profiles within each geographic division and reference year by baseline cell.

Booklet	X1B	X2B	X3B	X4B	X5B	X6B	X1C	X2C	X3C	X4C	X5C	X6C
1	1	2	1	2	2	6	2	1	1	1	2	5
1	2	1	2	3	3	3	3	1	2	3	1	1
1	3	3	3	1	2	2	3	3	2	2	2	1
2	1	1	2	2	2	4	1	1	3	1	2	5

[1] EPA assigned each profile to an independent subset, or "block" of profiles. Blocking reduces the number of profiles each respondent sees, thus reducing respondent burden.

2	2	2	1	3	1	5	1	2	1	3	2	6
2	3	3	3	2	3	3	3	3	2	3	2	2
3	2	2	2	1	2	2	1	2	1	3	2	5
3	2	3	1	2	3	6	1	1	1	2	3	2
3	3	1	3	3	1	1	2	1	3	3	3	2
4	1	1	2	2	3	3	1	3	3	2	1	4
4	2	3	3	3	1	1	2	2	3	3	2	1
4	3	1	2	2	3	5	3	1	2	1	2	2
5	1	1	3	1	1	4	1	2	1	1	2	4
5	2	3	3	2	3	6	1	3	3	2	1	2
5	3	2	3	2	3	3	3	3	2	3	2	3
6	1	1	2	1	3	3	2	2	3	1	3	4
6	2	1	1	2	2	3	2	3	1	2	1	5
6	3	1	3	3	1	6	3	2	3	2	1	5
7	1	1	2	2	3	2	2	1	3	2	1	4
7	1	3	3	1	2	1	1	2	3	3	2	2
7	2	2	1	3	1	3	3	2	1	1	1	1
8	2	1	1	1	3	2	2	1	3	1	2	4
8	2	2	2	2	3	1	2	2	2	3	2	1
8	3	3	2	3	1	4	1	3	1	3	1	3
9	1	1	2	3	3	5	1	3	2	3	2	6
9	1	2	2	1	1	6	1	1	1	1	1	3
9	3	1	1	2	2	1	2	1	3	2	2	3
10	1	2	1	3	1	1	1	3	1	3	2	3
10	1	3	3	2	3	3	2	3	3	3	3	5
10	3	1	2	1	2	6	1	1	2	1	1	3
11	2	1	3	1	3	1	2	2	2	1	3	1
11	2	3	1	2	2	6	1	3	2	2	2	5
11	3	2	3	1	3	3	3	2	3	2	2	4
12	2	2	1	1	1	3	2	1	1	3	1	4
12	3	1	2	2	2	1	2	2	2	2	2	1
12	3	3	2	3	3	2	3	2	3	3	3	1
13	1	3	1	1	3	1	3	3	1	2	3	4
13	2	2	2	1	3	6	2	2	2	2	1	4
13	2	3	1	3	2	5	3	1	1	3	2	3
14	1	3	1	2	3	1	2	2	1	2	3	2
14	2	1	2	1	2	4	2	1	2	3	1	6
14	3	2	2	2	1	6	3	1	3	2	1	5
15	1	2	1	3	3	5	3	2	1	1	3	4
15	2	3	2	3	2	3	2	3	3	2	2	4
15	3	1	1	2	1	6	3	1	1	1	2	6
16	1	2	3	1	1	5	1	3	2	1	1	6
16	2	2	3	3	3	6	3	2	3	1	3	5
16	3	1	1	2	3	1	3	3	1	1	3	2
17	1	1	3	2	2	5	1	2	3	2	1	6
17	2	3	2	1	3	6	1	3	3	1	3	6
17	3	1	1	3	1	6	2	1	1	1	1	4
18	1	2	2	2	3	3	1	1	1	2	3	1

18	2	3	3	1	2	2	2	2	3	2	2	3
18	3	2	2	2	1	2	3	2	3	3	1	3
19	1	1	3	3	2	1	3	1	3	3	1	2
19	1	3	2	3	1	4	1	2	2	2	1	2
19	3	2	2	1	3	5	3	3	1	1	3	6
20	1	2	2	3	2	2	1	2	1	3	3	4
20	2	3	2	2	1	5	2	2	1	2	1	2
20	3	3	3	3	3	4	2	3	3	3	1	2
21	2	3	2	1	1	1	1	3	3	1	1	1
21	3	2	3	2	2	6	3	2	3	1	1	3
21	3	3	1	3	3	4	3	2	1	3	2	2
22	1	3	1	1	1	1	2	3	1	1	2	3
22	2	1	3	3	3	6	1	1	3	1	3	2
22	2	3	2	3	3	5	1	2	2	3	3	4
23	2	1	2	2	1	5	2	1	1	1	1	2
23	3	1	1	3	3	5	1	1	2	3	3	4
23	3	2	2	1	2	4	3	3	2	1	1	3
24	1	2	3	1	2	5	1	1	3	3	2	6
24	1	3	1	2	1	2	3	3	1	2	2	4
24	3	3	2	1	1	5	3	3	1	1	2	5

Table A15-3: Attribute Levels Included in Each Survey Version.

Attribute	**Attribute Levels**						
	Baseline	**1**	**2**	**3**	**4**	**5**	**6**
	Declining Baseline						
Water Clarity (feet)	2	3	3.5	4.5	-	-	-
Adult Striped Bass (millions)	21	24	30	36	-	-	-
Adult Blue Crab (millions)	235	250	285	328	-	-	-
Oysters (tons)	2,800	3,300	5,500	10,000	-	-	-
Low Algae Level Lakes	2,300	2,900	3,300	3,850	-	-	-
Annual Household Cost	-	$20	$40	$60	$180	$250	$500
	Constant Baseline						
Water Clarity (feet)	3	3	3.5	4.5	-	-	-
Adult Striped Bass (millions)	24	24	30	36	-	-	-
Adult Blue Crab (millions)	250	250	285	328	-	-	-
Oysters (tons)	3,300	3,300	5,500	10,000	-	-	-
Low Algae Level Lakes	2,900	2,900	3,300	3,850	-	-	-
Annual Household Cost	-	$20	$40	$60	$180	$250	$500
	Improving Baseline						
Water Clarity (feet)	3.3	3.3	3.5	4.5	-	-	-
Adult Striped Bass (millions)	26	26	30	35	-	-	-

Adult Blue Crab (millions)	260	260	312	340	-	-	-
Oysters (tons)	4,300	4,300	5,500	10,000	-	-	-
Low Algae Level Lakes	3,100	3,350	3,600	3,850	-	-	-
Annual Household Cost	-	$20	$40	$60	$180	$250	$500

Realized Sample Sizes for Maximum Acceptable Sampling Error

The goal of the choice experiment is to estimate regression coefficients from mixed or conditional logit models that may be used to estimate willingness to pay for multi-attribute policy alternatives, or the likelihood of choosing a given multi-attribute alternative, following standard random utility modeling procedures (Haab and McConnell 2002). Hence, the sample size requirements are determined by the accuracy of the parameter estimates in the WTP models.

The resulting sample design will be a single stage stratified sample. No clustering (multiple stages of selection) will be necessary. Unequal probabilities of selection will result in different geographic divisions defined in Part B, Section 2 "Survey design", and lead to varying sampling weights, as demonstrated in Table A15-4 (assuming that the design contains 9 cells). Due to these varying weights, under assumptions of constant response rate and fixed sample size, the expected design effect due to differential baseline weights is 1.75. The realized design effect will likely be higher due to extra variability of weights within cells due to non-response adjustments.

Table A15-4. Sample size and accuracy projections.

Geographic division	Population size	Expected sample size	Expected weights	Standard error, 50% incidence	Standard error, 10% incidence
Bay States	5,479,176	1,728	3,171	0.017	0.010
Watershed	13,442,787	1,728	7,779	0.017	0.010
Other East Coast	25,431,478	1,728	14,717	0.017	0.010
Overall	44,353,441	5,184	8,556	0.011	0.007

Source: The household population size for each region was obtained from U.S. Census Bureau (2012). 2010 Census Summary File 1. Retrieved May 31, 2012 from http://factfinder2.census.gov/.

The maximum acceptable sampling error for predicting response probabilities (the likelihood of choosing a given alternative) in the present case is ±10%, assuming a true response probability of 50% associated with a utility indifference point. Given the survey population size, this level of precision requires a minimum sample size of approximately 96 observations. The number of observations (completed surveys) required to obtain large sample properties for the choice experiment design provide more than sufficient observations to obtain this required precision for population parameters.

Projected sample sizes given the potential non-response

Survey non-response is a common phenomenon. The sample design must be proactive and account for the potential non-response. Based on recent experience with surveys of similar nature, EPA expects the response rate for the Chesapeake Bay survey to be close to 30%. Additionally, the expected eligibility rate for a mail survey is 92%, and accounts for vacant, seasonal, non-existent, and otherwise ineligible units. The projected number of required mailings is given in Table A15-5 for different scenarios (response rates of 20% and 30%) and different sample size determination methods (expected number of mailings vs. the number of mailings that ensures 90% probability of reaching the cell target sample size).

Table A15-5. Required sample size.

Target cell size: *n*=288	*r*=20% response rate		*r*=30% response rate	
	Mean projection, *n/r*	90% prob to achieve cell size	Mean projection, *n/r*	90% prob to achieve cell size
Required cell size	1,565	1,672	1,043	1,111
District of Columbia	458	488	304	324
Maryland	3,696	3,948	2,462	2,622
Virginia	5,238	5,596	3,490	3,714
Delaware	240	256	160	170
New York	5,112	5,462	3,406	3,626
Pennsylvania	3,506	3,746	2,336	2,486
West Virginia	534	570	356	378
Connecticut	506	540	340	360

Florida	2,740	2,930	1,826	1,946
Georgia	1,324	1,414	882	940
Maine	206	220	138	146
Massachusetts	940	1,004	626	668
New Hampshire	192	204	128	136
New Jersey	1,186	1,268	790	842
North Carolina	1,382	1,478	922	982
Rhode Island	152	164	102	108
South Carolina	666	710	444	472
Vermont	94	102	64	68
Total:	28,172	30,100	18,776	19,988

The sample size required for 90% probability of achieving the cell size is computed as the 90-th percentile of the negative binomial distribution with success probability equal to the response rate and the number of successes equal to the target cell size.