End Commercial Tobacco Campaign: Parks Outdoor Secondhand Smoke Wave 1 Data Analysis Guidance

Issued: April 2022

Overall Guidelines

1) The Local Lead Agency (LLA) is responsible for the dataset and must keep a record of anyone requesting the data and with whom the data are shared (see the sample Data Request Tracking Form or contact the Tobacco Control Evaluation Center [TCEC] for assistance).
2) Local partner programs or others interested in obtaining the End Commercial Tobacco Campaign (ECTC) Parks Outdoor Secondhand Smoke dataset need to contact the LLA to request the data in writing and sign a form agreeing to data sharing and use guidelines.
    a. See sample Data Request Form or contact TCEC for assistance.
3) The following resources will be provided by TCEC along with the dataset to assist with LLA-level analysis:
    a. Codebook ("Parks Codebook" tab in the Excel file of each dataset) and the online survey
    b. This data analysis guidance document
    c. Data cleaning documentation ("Data Cleaning Summary" tab in the Excel file of each dataset)
    d. Training manual for question wording, explanations for each question, and online survey instructions saved as a PDF
4) Contact TCEC at tobaccoeval@ucdavis.edu with questions about the End Commercial Tobacco Campaign data analysis and reporting.

Sampling and Weights

1) **Sampling method**:
    a. The California Tobacco Control Program (CTCP) drew a random probability sample of parks which LLAs were required to survey. The list of parks was based on a December 2021 list from the California Protected Areas Database (CPAD) and from city, county, and community service district websites. See the TCEC Website for the complete sampling plan.
2) **Weights:**
    a. CTCP will calculate and provide survey weights to account for non-response 1-2 months after the completion of data collection by all LLAs.
    b. Weighted analyses are suggested, as they allow for results to more accurately represent the LLA community surveyed.
    c. However, weighted analyses are difficult in Excel. Thus, unweighted analyses are acceptable with the understanding that relying on

unweighted data will change the interpretation and representativeness of results.

d. Confidence intervals should be generated regardless of whether a census was attempted. Confidence intervals are needed due to potential non-response (e.g., data collector unable to survey a park), sampling frame errors (e.g., missing park not in the sampling frame), and measurement error (e.g., data collector accidentally recording the wrong data). These can increase the variability of the estimates and introduce bias. In addition, the longitudinal nature of ECTC requires confidence intervals.
    i. Unweighted Analyses in **Excel**, **SPSS**, or **SAS**: Use a higher 99% confidence level and specify in report or footnote.
    ii. Weighted Analyses or a Census: Use the default 95% confidence levels.
        1. **In SPSS**, account for the calculated sampling weight using the following point and click instructions before conducting descriptive analyses: Data → Weight Cases → Weight cases by.

           Example:

           ```
           WEIGHT BY <var name>.
           ONEWAY var1 BY var2
             /STATISTICS DESCRIPTIVES
             /MISSING ANALYSIS
             /CRITERIA=CILEVEL(0.95).
           ```
        2. **In SAS**, use "PROC SURVEYFREQ" or "PROC SURVEYMEANS" and specify the weighting variable.

           Example:

           ```
           PROC SURVEYFREQ data = <dataset>;
           table var1*var2 / row cl;
           weight <var name>;
           RUN;
           ```
        3. If the LLA was able to complete a census of parks, they may use **Excel** to calculate a 95% confidence interval. Completing a census means the LLA was able to observe and collect data at all parks.

Data Analysis and Reporting

1) LLAs should consider their target audience and how they want to use the results of their data when developing the data analysis plan and framing any reports of results. Different variables, different sub-analyses, and different language may be appropriate for different audiences.
    a. For example, the LLA may want to consider whether they are educating policymakers interested in jurisdiction-wide results on tobacco product waste or if they are engaging with neighborhood groups that work with children and may be interested in

neighborhood-level data including sub-analysis for evidence of smoking near playgrounds.
    b. Sub-groups of parks may be combined and analyzed to assess parks with similar demographic or geographic characteristics (e.g., neighborhood socioeconomic status, urban/suburban/rural location) as deemed useful by the LLA.
    c. Use the example reporting language provided in this document to frame the presentation of findings.
2) TCEC will provide guidance on analyzing data using Excel. Example SAS code for recoding variables and answering some evaluation questions are presented in Appendices 1 and 2. Assistance with other statistical programs is available upon request.
3) LLA-level analyses will typically aim to assess Tobacco Product Waste (TPW) and Observed Active Smoking (OAS) on average across all jurisdictions and for various jurisdictions, parks, and specific park amenities (e.g., playgrounds).
    a. Use the codebook, training manual, and this data analysis guidance document to understand which questions to use for results that may be of interest. Examples of key questions for analysis are included in this document, but LLAs may decide to analyze other variables as well.
4) LLAs should use caution when reporting information on individual parks, particularly private parks.

Variables

1) Variable names are listed in the Codebook ("Parks Codebook" tab in the Excel file of each dataset).
2) **Predictor Variables:**
    a. Park ID (PQID) and Jurisdiction (Community)
        i. Categorical Variables
        ii. LLAs should to data checks on the Jurisdiction (Community) to ensure that the variable information is correct.
3) **Outcome Variables:**
    a. The main outcome variables to include in the descriptive analysis are TPW and Observed Active Smoking (OAS). LLAs may also be interested in analyzing No Smoking or Vaping Signage.
    b. Measures are available for each amenity Type (e.g., playground, picnic, restroom, sports field, athletic court, parking lot).
        i. LLAs can assess these measures separately for each amenity or they can compute an average across all amenities (see Appendix 1).
            a. Numeric outcome measures (e.g., TPW) might be skewed due to outliers. LLAs should conduct data checks of skew and/or kurtosis for the presence of outliers. If skew/kurtosis is high (±2),

it is suggested that median values (instead of means) are reported.

ii. LLAs may wish to recode these variables into a categorical variable prior to analysis (see Appendix 1).

c. Tobacco Product Waste (TPW):
   i. *Survey Question*: How many pieces of TPW are in the observation area?
   ii. *Numeric Measure*: Open-ended response scale assessing count (ranging from 0 to 99)
   iii. *Categorical Recode*: Recode the numeric measure of TPW to binary/categorical variable where:
      a. 0 = 0: *no TPW*
      b. 1 thru 99= 1: *yes, TPW observed*

d. Observed Active Smoking (OAS):
   i. *Survey Question*: During your observation, how many times did you see or smell tobacco or marijuana smoke or vapor?
   ii. *Numeric Measure*: Response options ranging from 0 to 5 or more
   iii. *Categorical Recode*: Recode the numeric measure of OAS to binary/categorical variable where:
      a. 0 = 0: *no Observed Active Smoking*
      b. 1 thru 5 = 1: *yes, Observed Active Smoking*

e. "No Smoking/No Vaping" Signage:
   i. *Survey Question*: Is there a "No smoking/No vaping" sign nearby?
   ii. *Categorical Measure*: Response options were yes/no.

Interpretation and Write-Up Guidance

1) For each analysis, present an estimate (estimated average or estimated percentage) and corresponding confidence interval.
2) Results of analyses that assess TPW should be reported per 100 square feet.
3) If the LLA used weighted data or completed a census and is assessing TPW by jurisdiction, then 95% confidence intervals should be calculated, and data can be reported as:
   a. **In 2022 in Y Jurisdiction, approximately 5.0 (95% CI: 3.1, 5.2) pieces of tobacco product waste per 100 square feet were observed in each park on average.**
4) If the LLA used unweighted data and is assessing TPW by jurisdiction then 99% confidence intervals should be calculated, and data can be reported as:
   a. **In 2022 in the parks surveyed in Y Jurisdiction, approximately 5.0 (99% CI: 4.9, 5.1) pieces of tobacco product waste per 100 square feet were observed in each park on average.**
5) If the LLA used weighted data or completed a census and is assessing the percentage of parks that had any TPW (> 0) across all jurisdictions, 95% confidence intervals should be calculated, and data can be reported as:

a. **In 2022 across X Jurisdiction, tobacco product wase was observed in approximately 25.0% (95% CI: 23.1%, 26.9%) of parks.**

Example Evaluation Questions with Data Analysis Suggestions

Predictor Variable(s)	Outcome Variable(s)	Outcome Variable Type	Reported Estimates
What is the average amount of TPW per jurisdiction?			
Jurisdiction (Community)	TPW (average across all amenities, e.g., mean of PAPGTPW, PAPNTPW, PARRTPW, etc.)[a]	Numeric	Average/mean[b] and confidence interval of TPW for each jurisdiction
What is the average amount of TPW by park and jurisdiction?			
Park ID (PQID) *and* Jurisdiction (Community)	TPW (average across all amenities, e.g., mean of PAPGTPW, PAPNTPW, PARRTPW, etc.)[a]	Numeric	Average/mean[b] and confidence interval of TPW for each park and for each jurisdiction
What percentage of parks had any TPW per jurisdiction?			
Jurisdiction (Community)	TPW (average across all amenities, e.g., mean of PAPGTPW, PAPNTPW, PARRTPW, etc.) Recoded as No/Yes TPW (TPW_cat)[a]	Categorical	Percentage (%) and confidence interval estimate of parks with TPW (1 or Yes TPW) for each jurisdiction
Which areas of the parks had the most TPW on average by jurisdiction?			
Jurisdiction (Community)	TPW at each separate amenity (PAPGTPW, PAPNTPW, PARRTPW, PASFTPW, PAACTPW, PAOATPW, PAPLTPW)[a]	Numeric	Average/mean[b] and confidence interval of TPW for each amenity type and for each jurisdiction
What was the average amount of TPW at playgrounds by jurisdiction?			
Jurisdiction (Community)	TPW at the Playground Amenity (PAPGTPW)	Numeric	Average/mean[b] and confidence interval of TPW at playgrounds for each jurisdiction
What percentage of parks had any Observed Active Smoking across by jurisdiction?			
Jurisdiction (Community)	Observed Active Smoking (average across all amenities, e.g., mean of PAPGM, PAPNM, PARRM, etc.) Recoded as No/Yes OAS (OAS_cat)[a]	Categorical	Percentage (%) and confidence interval estimate of parks with any Observed Active Smoking (1 or Yes OAS) for each jurisdiction

Predictor Variable(s)	Outcome Variable(s)	Outcome Variable Type	Reported Estimates
Which areas of the park had the most Observed Active Smoking on average by jurisdiction?			
Jurisdiction (Community)	Observed Active Smoking at each separate amenity (PAPGM, PAPNM, PARRM, PASFM, PAACM, PAOAM, PAPLM)[a]	Numeric	Average/mean[b] and confidence interval of Observed Active Smoking for each amenity type and for each jurisdiction
What percentage of parks had any No Smoking or Vaping Signage by jurisdiction?			
Jurisdiction (Community)	No Smoking or Vaping Signage (average across all amenities, e.g., mean of PAPGS, PAPNS, PARRS, etc.) Recoded as No/Yes Signage (Signage_cat)[a]	Categorical	Percentage (%) and confidence interval estimate of parks with any No Smoking or Vaping Signage (1 or Yes Signage) for each jurisdiction
Which areas of the parks were most likely to have No Smoking or Vaping Signage by jurisdiction?			
Jurisdiction (Community)	No Smoking or Vaping Signage at each separate amenity (PAPGS, PAPNS, PARRS, PASFS, PAACS, PAOAS, PAPLS) Recoded as No/Yes Signage (Signage_cat)[a]	Categorical	Percentage (%) and confidence interval estimate of No Smoking or Vaping Signage (1 or Yes Signage) for each amenity type and for each jurisdiction

Note. Example dataset variable names are listed in blue.
a. See Appendix 1.
b. If skew/kurtosis is high (±2), median values (instead of mean values) may be reported.

Appendix 1

Detailed Examples of SAS Variable Re-Coding

Tobacco Product Waste (TPW):

Compute mean/average score of TPW across all amenities:

```
TPW=mean(PAPGTPW, PAPNTPW, PARRTPW, PASFTPW, PAACTPW, PAOATPW,
PAPLTPW);
run;
```

Categorical Recode:

Option 1:

```
if TPW=0 then TPW_cat='No TPW';
else if TPW in (1:99) then TPW_cat='Yes TPW';
```

Option 2:

```
if TPW=0 then TPW_cat=0;
else if TPW in (1:99) then TPW_cat=1;
proc format;
value TPWformat
      0 = ' 0: No TPW '
      1 = ' 1: Yes TPW ';
```

Observed Active Smoking (OAS):

Compute mean/average score of OAS across all amenities:

```
OAS=mean(PAPGM, PAPNM, PARRM, PASFM, PAACM, PAOAM, PAPLM);
run;
```

Categorical Recode:

Option 1:

```
if OAS=0 then OAS_cat='No OAS';
else if OAS in (1:5) then OAS_cat='Yes OAS';
```

Option 2:

```
if OAS=0 then OAS_cat=0;
else if OAS in (1:5) then OAS_cat=1;
proc format;
value OASformat
      0 = ' 0: No OAS '
      1 = ' 1: Yes OAS ';
```

No Smoking or Vaping Signage:

Compute mean/average score of Signage across all amenities:

```
Signage=mean(PAPGS, PAPNS, PARRS, PASFS, PAACS, PAOAS, PAPLS);
run;
```

Categorical Recode:

Option 1:

```
if Signage =0 then Signage_cat='No Signage';
else if Signage in (1:7) then Signage_cat='Yes Signage';
```

Option 2:

```
if Signage =0 then Signage_cat=0;
else if Signage in (1:7) then Signage_cat=1;
proc format;
value Signageformat
      0 = ' 0: No Signage '
      1 = ' 1: Yes Signage ';
```

Appendix 2

Detailed Examples of SAS Analysis Code for Obtaining Overall Estimate and Corresponding Confidence Interval

Example 1: Weighted estimated average/mean value and corresponding confidence interval for numeric outcome measures

Use: PROC SURVEYMEANS

SAS Code: TPW across all Parks:

```
DATA <dataset>; set <original dataset>;
TPW=mean(PAPGTPW, PAPNTPW, PARRTPW, PASFTPW, PAACTPW,
PAOATPW, PAPLTPW);
RUN;

PROC SURVEYMEANS data = <dataset name> plots=none;
var TPW;
weight <var name>;
RUN;
```

SAS Results Output:

Statistics					
Variable	N	Mean	Std Error of Mean	95% CL for Mean	
TPW	23	3.121298	0.713210	1.64219013	4.60040507

Interpretation:

In 2022 in Jurisdiction X, an average of 3.1 (95% CI: 1.6, 4.6) pieces of tobacco product waste per 100 square feet was observed.

Example 2: Weighted estimated percentage and corresponding confidence interval for categorical outcome measures

Use: PROC SURVEYFREQ

SAS Code: Observed Active Smoking (OAS) in the Playground Amenity:

```
DATA <dataset>; set <original dataset>;
if PAPGM=0 then PAPGM_cat=0;
else if PAPGM in (1:5) then PAPGM_cat=1;
RUN;

PROC FORMAT;
value PAPGMformat
      0 = ' 0: No OAS '
      1 = ' 1: Yes OAS ';
RUN;

PROC SURVEYFREQ data = <dataset>;
weight <var name>;
table PAPGM_cat / row cl;
format PAPGM_cat PAPGformat.;
RUN;
```

Table of PAPGM_cat							
PAPGM_cat	Frequency	Weighted Frequency	Std Err of Wgt Freq	Percent	Std Err of Percent	95% Confidence Limits for Percent	
0: No OAS	78	166.27000	15.97079	46.3264	4.2113	38.0115	54.6413
1: Yes OAS	88	192.64000	16.50084	53.6736	4.2113	45.3587	61.9885
Total	166	358.91000	11.79399	100.0000			

Interpretation:
In 2022 in Jurisdiction A, active smoking was observed in 53.7% (95% CI: 45.4%, 62.0%) of playgrounds.

Example 3: Unweighted estimated average/mean value and corresponding confidence interval for numeric outcome measures for multiple jurisdictions within the same county.
Use: PROC SURVEYMEANS
SAS Code: TPW by Jurisdiction Across all Amenities:

```
DATA <dataset>; set <original dataset>;
TPW=mean(PAPGTPW, PAPNTPW, PARRTPW, PASFTPW, PAACTPW,
PAOATPW, PAPLTPW);
RUN;

PROC SURVEYMEANS data = <dataset name> alpha=0.01
plots=none;
domain Jurisdiction;
var TPW;
RUN;
```

SAS Results Output:

Statistics for Jurisdiction Domains						
Jurisdiction	Variable	N	Mean	Std Error of Mean	99% CL for Mean	
Jurisdiction 1	TPW	13	2.461538	0.492632	1.07292897	3.85014796
Jurisdiction 2	TPW	10	3.200000	1.070429	0.18272185	6.21727815

Interpretation:
In 2022 across parks surveyed in Jurisdiction 1, an average of 2.5 (99% CI: 1.1, 3.9) pieces of tobacco product waste per 100 square feet was observed. In Jurisdiction 2, across all surveyed parks, an average of 3.2 (99% CI: 0.2, 6.2) pieces of tobacco product waste per 100 square feet was observed.