Bonneville Power Administration

2020-2021 Evaluation Plan

Submitted by Evergreen Economics

In partnership with Apex Analytics, LLC and SBW Consulting

January 8, 2020

1 OVERVIEW

The Bonneville Power Administration (BPA), along with its public power utility partners, acquires savings from a portfolio of energy efficiency programs and measures. Currently, the portfolio includes the following measures and savings estimation techniques:

- Unit Energy Savings (UES) measures utilizing a constant savings value for each measure application;
- Custom measures, requiring calculation of savings for each project; and
- Calculator measures with a standardized savings estimation algorithm and project-specific parameter values (typically lighting).

This document provides a plan that builds off of BPA's recent impact evaluation efforts. Specifically, this plan describes the activities and approaches the Evergreen Economics team (which includes Apex Analytics, SBW Consulting, and Demand Side Analytics) will undertake in fiscal years 2020 and 2021 (FY2020-21).

The following sections provide the background, context, and objectives of the FY2020-21 impact evaluation activities.

1.1 OVERVIEW OF BPA APPROACH TO IMPACT EVALUATION

Over time, BPA and the Regional Technical Forum (RTF) have developed a series of documents to provide guidance on how to estimate savings. Portions of these documents provide guidance on how to estimate savings from the projects that comprise the UES portfolio.

- **RTF Guidelines:**[1] The RTF uses guidelines to judge the quality and reliability of the savings estimates, costs, benefits, and lifetimes for all types of efficiency measures. The RTF will provide guidance on delivery verification for UES and Standard Protocols and specifies that 80 percent of the portfolio of savings should be evaluated every three years.
- **RTF Delivery Verification Requirements:** For each UES measure, the RTF identifies key data that need to be collected (or checked) to ensure reliability of the RTF savings estimate. These requirements include detailed checklists and updated measure specifications.
- **BPA Quality System Strategy & Implementation (QSSI):** QSSI presents a framework for establishing BPA's system that is used to ensure high-quality programmatic energy savings or "quality system." This quality system framework focuses on programmatic energy savings. It includes Standards; Planning Policies; Oversight Policies; Impact and Process Evaluation Policies; and Savings Policies for Custom Projects, Calculators, and Unit Energy Savings Measures. Regarding impact evaluation, the QSSI policies state that BPA conducts impact

[1] Regional Technical Forum, 2018 Full Operative Guidelines: https://nwcouncil.box.com/v/2018RTFOperativeGuidelines. Guidelines Section 5.1.2 defines that evaluation is required for large programs that account for more than 10 percent of portfolio savings that have not been evaluated in the previous three years, and programs that have substantially changed or have uncertain savings. The sum of the savings that are not evaluated should not exceed 20 percent of portfolio savings.

evaluation on 80 percent of its portfolio every four years, which aligns the duration with two-year rate periods.

- **BPA Implementation Manual:**[2] The manual, together with the customer's Energy Conservation Agreement (ECA) and specifications in BPA's energy efficiency reporting system, provide the implementation requirements for projects reported to BPA.
- **BPA M&V Protocols:** BPA M&V Protocols[3] direct custom project measurement and verification (M&V) activities to support reliable savings estimates. Protocols include the M&V Protocol Selection Guide; reference guides for sampling, regression, and glossary; and protocols on metering, indexing, engineering calculations with verification, energy modeling, and existing building commissioning. To support M&V of strategic energy management projects, M&V Protocols also include the Monitoring, Tracking and Reporting (MT&R) Reference Guide.[4]

1.2 EVALUATION COVERAGE

From 2012 through 2019, BPA conducted evaluation on 89 percent of the portfolio of savings.[5] Table 1 includes impact evaluations conducted since 2012. Consistent with BPA's QSSI policies, evaluation coverage effectively expires after four years. Therefore, Table 1 includes the status of coverage: expired (more than four years old), expires in 2021 (2017 evaluation) and current (less than four years).

Table 1: Historic Evaluations

Study Name	*Evaluation Coverage*	*Most Recent Year of Savings*	*Study Completion Date*	*Sector*	*Measures*
2019 Residential HVAC Impact Evaluation	Current / Covered	2018	TBD	Residential	Heat pumps: air-source, variable speed and ductless. Prescriptive duct sealing.
Impact Evaluation of Ductless Heat Pumps and Prescriptive Duct Sealing	Expired	2016	9/1/2018	Residential	Duct sealing, ductless heat pumps

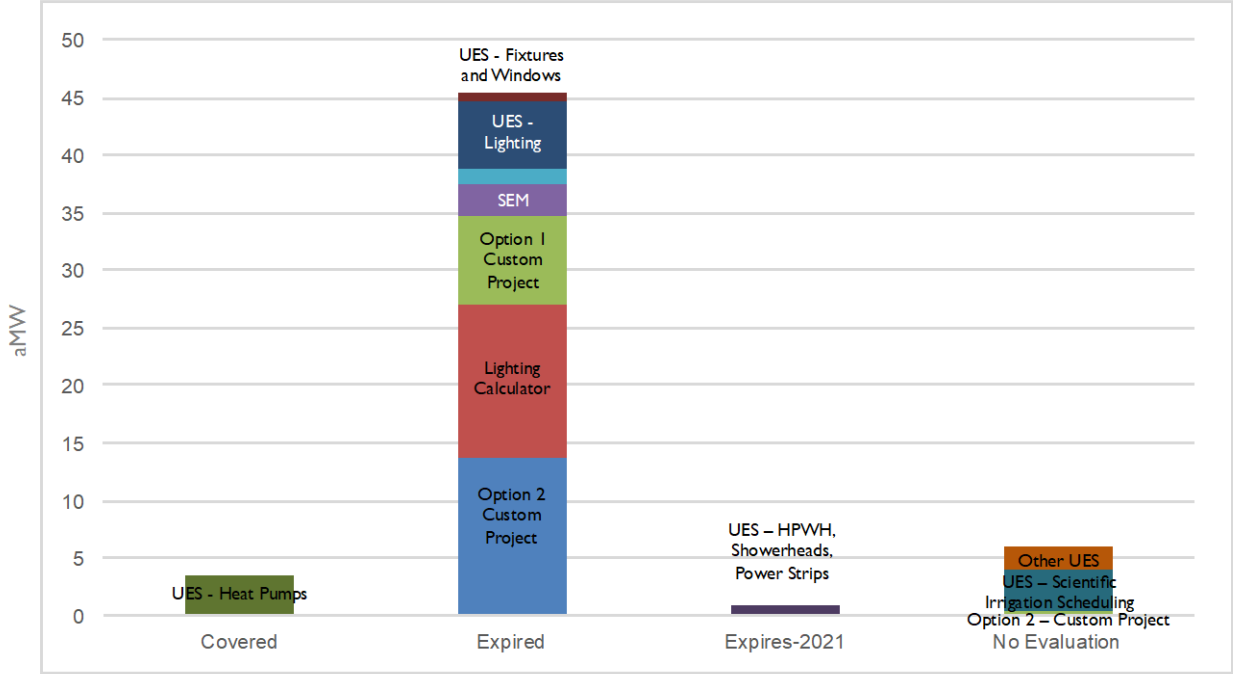
[2] Bonneville Power Administration. 2019. *2020-2021 Implementation Manual*. https://www.bpa.gov/EE/Policy/IManual/Documents/2020-2021_IM_Updated_9-17-19.pdf
[3] https://www.bpa.gov/EE/Policy/IManual/Pages/IM-Document-Library.aspx
[4] https://www.bpa.gov/EE/Policy/IManual/Documents/MTR-Reference-Guide-Rev7.pdf
[5] Based on FY2018 reported savings.

Study Name	Evaluation Coverage	Most Recent Year of Savings	Study Completion Date	Sector	Measures
Billing Analysis of Select Weatherization & HVAC Measures	Expired	2015	3/1/2018	Residential	Insulation & Windows
2017 Delivery Verification Results and Findings	Expires in 2021	2017	5/1/2018	Residential	Heat Pump Water Heaters, De-energization, BPA Green Motors, Power strips, showerheads
Impact Evaluation Site-Specific Portfolio	Expired	2013	11/1/2015	Commercial, Industrial, and Agriculture	Custom and Lighting Projects of Option 1, Option 2 and Federal.
Industrial Strategic Energy Management Impact Evaluation	Expired	2014	2/1/2017	Industrial	Energy Management
Residential Lighting UES Impact Evaluation	Expired	2015	2016	Residential	Lamps and Fixtures

Figure 1 shows the evaluation coverage status for savings categories (e.g., custom measures, UES measures, strategic energy management). The majority of historic evaluation coverage has expired, due in large part to the expiration of Site-Specific Savings impact evaluation. This evaluation included FY2012-13 savings of non-residential lighting and custom projects for Option 1 and Option 2 utilities.[6]

[6] Option 1 utilities use BPA M&V services to support their commercial and agricultural custom projects. Option 2 utilities use their internal staff engineering resources to complete M&V on their commercial and agricultural custom projects. Energy Smart Industrial technical services are used by both Option 1 and Option 2 utilities.

Figure 1: Historic Evaluation Coverage



Approximately 7 percent of the FY2018 portfolio savings is considered covered by recent evaluation. This represents the current assessment of the Performance Tested Comfort Systems (PTCS) program, specifically including air-source and variable-speed heat pumps, ductless heat pumps, and prescriptive duct sealing. The assessment includes analysis of customer billing data pre- and post-measure installation, a customer survey, and program data such as characteristics of the home and baseline measures.

1.3 EVALUATION PLANNING ACTIVITIES

To develop this FY2020-21 evaluation plan, the Evergreen team first conducted a detailed assessment of the portfolio to understand evaluation coverage, priorities, and opportunities. The team linked the data sources shown in Table 2.

Table 2: Data Sources Used for Portfolio Assessment

Source	*Information*
BPA reporting system data[7]	Historic energy savings achievements since 2012 for all measures by utility.
UES Measure List	Detailed measure characteristics, such as lifetime and baseline.
RTF Measure Status	Details of current measure status from the RTF, including whether a measure is Proven or Planning.

[7] Business Operations Report provided by BPA in July 2019.

Source	Information
Historic evaluations	Timing of historic evaluations, by measure, were linked to allow for estimation of savings coverage of the current portfolio.
Engineering and program priorities	BPA engineers and program staff assessed measures and identified their perspective of priorities.
Other research activities	Research activities of BPA's Momentum Savings, Northwest Energy Efficiency Alliance, and the RTF were collected to identify opportunities and potential overlaps.

Based on this assessment, the Evergreen team conducted an evaluation seminar with internal BPA staff to share the results and discuss evaluation opportunities. Appendix E has detailed tables of UES measure considerations and reasons for inclusion/exclusion.

1.4 PROPOSED 2020-2021 ACTIVITIES

Table 3 on the next page summarizes the FY2020-21 evaluation plan. Using the data assessment, the Evergreen team identified that the highest priority is to conduct evaluation on custom and commercial, industrial, and agriculture (C/I/Ag) lighting projects. These projects, which require M&V or calculators for savings estimation, represent approximately 65 percent of the portfolio and, as noted above, were most recently evaluated for FY2012-13 savings. Additionally, the FY2020-21 plan includes strategic energy management (SEM) evaluation, assessment of BPA Qualified Commercial HVAC measure savings, and evaluation of Clark Public Utilities' multifamily thermostat pilot.

Table 3: Summary of FY2020-21 Evaluation Plan

Evaluation Areas	Brief Summary	Savings % of Portfolio
Custom and C/I/Ag Lighting	Rolling, engineering-based evaluation.	65%
Strategic Energy Management	Persistence assessment to inform measure life and (if feasible) assessment of how capital measures affect SEM savings.	5%
BPA Qualified (BPAQ) Commercial HVAC	Billing analysis to support BPAQ measure assessment.	<1%
Clark Public Utilities thermostat pilot	Billing analysis to understand program savings.	NA

Table 4 below shows the timeline of planning, data collection, analysis, and reporting of the evaluation areas of this plan. The Residential HVAC evaluation that the Evergreen team began in FY2019 is also included.

Table 4: Timeline of Evaluation Activities

Evaluation Activity		FY2020 Q1	FY2020 Q2	FY2020 Q3	FY2020 Q4	FY2021 Q1	FY2021 Q2	FY2021 Q3	FY2021 Q4
Custom and C/I/Ag Lighting*	Option 1 Custom Industrial								
	Option 2 Custom Industrial								
	Option 1 Lighting								
	Option 2 Lighting								
SEM Evaluation									
BPAQ Commercial HVAC									
Clark Thermostat Pilot Evaluation									
(FY2019) Residential HVAC Evaluation									

** Non-Industrial Custom (Option 1 and Option 2) and Energy Smart Reserve Power are planned for FY2022-23. See Appendix A for more detail.*

The following sections provide an overview of the FY2020-21 proposed activities, and Appendices A-D provide details on each evaluation area.

1.4.1 OTHER EVALUATION CONSIDERATIONS

Across evaluation activities, we will consider the following factors:

- **Time-based value of energy efficiency:** Due to the increased importance of the time-based value of energy efficiency, the evaluation will include activities to support the assessment temporal savings, wherever possible. This is not a primary focus of the evaluation but will be a secondary consideration when it is cost-efficient to collect time-based information on energy efficiency measures.
- **Non-energy impacts and customer satisfaction:** If customers are contacted for impact evaluation purposes, there is an opportunity to collect non-energy impact (NEI) information, as well as assess customer satisfaction. BPA and the Evergreen team will collaborate to define protocols for these factors when impact evaluation surveys are conducted.

Additionally, the Evergreen team makes the following recommendations for BPA during the FY2020-21 timeframe:

- **Recommendations tracker:** As BPA conducts more evaluations, it is valuable to track historic recommendations and responses or progress toward those recommendations. Therefore, we recommend BPA develop a simple recommendations tracker that includes historic evaluation recommendations and progress toward those recommendations (or reasons for not implementing the recommendations).
- **Advise on the evaluability of new measures:** As needed, the Evergreen team will advise BPA on the evaluability of new measures.

- **Quarterly evaluation updates:** At the beginning of each quarter, BPA should provide an update on evaluation activities, which are uploaded to the BPA Evaluation website and included in customer announcements.
- **Stakeholder and communication protocols:** The Evergreen team will coordinate with BPA on developing stakeholder and communication protocols for evaluation work in general as well as individual evaluation projects.

1.4.2 FY2020-21 EVALUATION PLAN UPDATE

In the fourth quarter of 2020, the Evergreen team will update the data assessment described in Section 1.3 to include FY2019 and FY2020 data. This update could result in an adjustment of priorities for FY2021. In particular, the evaluation team will assess the following measures for increased savings achievements or change in status, which could lead to a need for additional evaluation:

- Commercial thermostats
- Commercial & Industrial (C&I) lighting controls
- Residential showerheads and thermostatic shutoff valves
- Power strips (residential and commercial sectors)
- Agriculture hardware and pumps
- Residential Low-E windows
- Commercial pumps and fans
- Commercial variable frequency drives (VFDs)
- Residential commissioning/controls/sizing efficient case research
- Low-rise multifamily sector research
- Understanding building shell for residential HVAC

2 EVALUATION METHODOLOGIES

This section describes the approaches, data collection efforts, and sample frame that the Evergreen team will use for the impact evaluation. These methodologies build on the guidelines set forth in the Quality System Strategy & Implementation (QSSI) document, Regional Technical Forum (RTF) Guidelines,[8] and the BPA Implementation Manual (IM).[9]

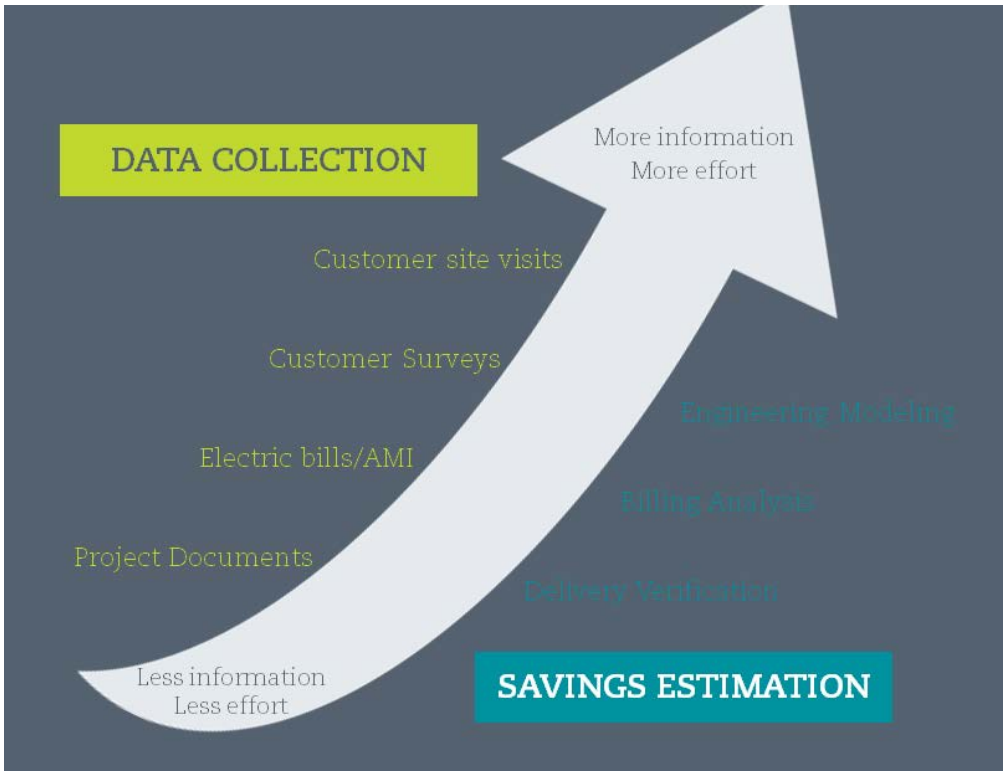
2.1 GENERAL APPROACHES

The evaluation team aims to select the best approach available to conduct the evaluation while balancing strategic considerations including a measure's status, contribution to savings, uncertainty in claimed savings, and programmatic importance.

Generally, there are multiple evaluation data sources and savings estimation methods, as shown in Figure 2 below. Data sources include project documents (typically files maintained by customer utilities with detailed project information), electricity bills, customer surveys (phone, mail, etc.), and customer site visits.

Delivery verification is typically used for RTF-proven UES measures, and billing analysis and engineering modeling are used for non-proven UES measures, standard protocol, and custom measures. Delivery verification can generally be completed via two approaches: review of project documentation or installation verification through end-user contact such as phone surveys or site visits. As such, delivery verification is lower effort for evaluation, and provides insight into total program savings by verifying quantity, but not about per-unit saving values. Assessing savings through billing analysis or engineering modeling, on the other hand, requires more evaluation effort, but yields greater insight into installed measure savings.

Figure 2: Data Collection and Savings Estimation Methods



2.2 2020-2021 DATA COLLECTION AND APPROACHES

In FY2020-21, the evaluation team plans to use engineering modeling for the Custom and C/I/Ag Lighting evaluation and SEM capital evaluation, and billing analysis for the BPA Qualified (BPAQ) Commercial HVAC assessment and the Clark Public Utilities' residential thermostat pilot evaluation. For all evaluation activities, the Evergreen team will collect and review project files and where needed,[8] the team will conduct customer surveys. Site visits may be required for some projects in the Custom and C/I/Ag Lighting evaluation and for the SEM capital evaluation, as shown in Table 5 below.

Table 5: Key Data Collection and Savings Estimation Method

	Data Collection Activities			*Saving Estimation Methods*	
	Billing Data	*Customer Surveys*	*Site Visits*	*Engineering Modeling*	*Billing Analysis*
Custom and C/I/Ag Lighting		✓	✓*	✓	

[8] Contacting customers is needed whenever the data provided does not support a reliable estimate of savings or additional information is required for evaluation that is not provided on project documentation. This is defined for each evaluation separately, and associated contact protocols will be followed if customer contact is required.

Strategic Energy Management	Persistence		✓	✓*		
	Capital		✓	✓*	✓	
BPAQ Commercial HVAC		✓	✓*			✓
Clark Thermostat Pilot		✓	✓*			✓

* As needed, lack of information in project documentation.

Primarily, the evaluation team will use project documentation (the documentation required per the Implementation Manual), billing data, phone surveys, and site visits to support the FY2020-21 evaluation activities. In order to function cost-effectively and efficiently, the evaluation will seek to leverage any and all data that are already collected from existing BPA and utility staff's data collection efforts.

- **Project documentation:** Project documentation may include data from Interim Solution 2.0 (IS2.0), files uploaded to BPA's Energy Efficiency Documents, data required in the Implementation Manual to be maintained by utilities, and any additional information collected by third party implementers or program staff. Following the contact protocols outlined in the appendices, the evaluation team will work with BPA staff and participating utilities to obtain utility customer documentation and files for each sampled measure, when necessary. If files are missing critical information, the evaluation team will work with BPA to determine if the additional information may be made available through a supplemental request.
- **Utility bills:** The evaluation team will request billing data to support the BPAQ Commercial HVAC assessment and the Clark Public Utilities residential thermostat pilot evaluation. In these instances, the evaluation will target a census of energy consumption data across the sampled utilities. To reduce the burden on utilities and streamline the billing data request process, the evaluation team will provide a data template at the time of sample notification, consistent with the 2019 template.
- **Customer surveys or site visit:** Where the evaluation approach includes customer contact (in the form of a phone survey or site visit) the evaluation team will follow contact protocols relevant for each evaluation activity. For FY2020-21, the Evergreen team will follow contact protocols (Appendix F).

Details on data collection and evaluation methods for each evaluation areas are in Appendix A-E.

2.3 SAMPLE DESIGN

This section provides a description of the general sampling strategy and the draft FY2020-21 sample design. BPA's QSSI policies have established a target for impact evaluation, striving for measure group-level evaluations to attain a relative error of 10 percent at the 90 percent confidence level, with a minimum acceptable level of 80/20.

The evaluation team-proposed sampling strategy targets a 90/10 confidence level and precision for the Custom and C/I/Ag Lighting evaluation, with approximately 80 measures sampled in FY2020 and FY2021.[9] The SEM evaluation will use a simple random sample approach, with approximately 15 sites for the persistence study and 10 sites for the capital projects.[10] Both the BPAQ Commercial HVAC assessment and the Clark Public Utilities residential thermostat pilot represent a census of sites for the statistical analysis. Table 6 also outlines the years that the sample will cover. The Custom and C/I/Ag Lighting and SEM capital evaluations will sample from the previous year of participation. The SEM persistence study will leverage sites from 2015-2019, while the BPAQ Commercial HVAC assessment and the Clark Public Utilities residential thermostat pilot evaluation will include approximately three years of program participants.

Table 6: FY2020-21 Evaluation Plan Sample Size

Area		*FY2020- FY2021 Sample Size*	*Technique*	*Sampled Years*
Custom and C/I/Ag Lighting		113[11]	Savings stratified random sample	One year prior
Strategic Energy Management	Persistence	~15	Simple random	2015-2019
	Capital	~10	Simple random	One year prior
BPAQ Commercial HVAC		~150	Census	2017-2019
Clark Public Utilities Thermostat Pilot		~1,800	Census	2018-2020

These draft sample sizes are based on FY2018 IS2.0 data pulled in July of 2019. Savings in the IS2.0 database is expected to change, and the evaluation team will revise the samples based on the final IS2.0 data prior to the start of each evaluation category.

2.3.1 UTILITY-SPECIFIC OVERSAMPLES

The draft sample design will most likely not support statistically reliable estimates of savings for utility-specific measure groups. However, additional studies can be added to the sample design that would support estimates for specific utilities.

If a utility is interested in conducting an oversample in its territory to gain statistical significance, the utility can contact the evaluation contractor. The evaluation

[9] Appendix A outlines additional sample of 80 proposed for FY2022-23.
[10] Although a simple random approach will be used, the evaluation team will first work with BPA staff to identify characteristics of sites that are included in the sampling. For example, the sample may only include sites with multiple years of participation and positive savings.
[11] Total four-year (FY2020-23) sample for this evaluation area is expected to be 165 projects. See Appendix A for more detail.

contractor will work with the utility to determine the sampling strategy for their study and the required confidence/precision. Participating utilities would have to separately contract with the evaluation team for the oversample.

BPA will fund the fixed costs associated with the impact evaluation (e.g., database development, sampling, evaluation protocols, training), and the utility requesting an oversample will fund the marginal costs of additional site-specific analysis costs (e.g., data collection and savings estimation). The utilities will also be responsible for any expenses associated with the preparation of utility-specific evaluation reports and presentations.

2.4 PROJECT MANAGEMENT

2.4.1 STAFFING

Evergreen Economics is the prime contractor responsible for the evaluation and will be reporting to Carrie Nelson, the Evaluation Lead, and Keshmira McVey, the Contracting Officer's Technical Representative (COTR) for BPA, who are supported by additional BPA staff Melissa Podeszwa, Phillip Kelvsen, and Michele Francisco. The organization of the evaluation team is designed to maximize project management and consistency, while maintaining a high level of quality control. Tami Rasmussen will act as the Project Manager, and Dr. Steve Grover is the Principal-in-Charge. The team also includes SBW Consulting, Apex Analytics, and Demand Side Analytics.

2.4.2 COORDINATION WITH BPA OVERSIGHT

BPA conducts reviews of utility-reported measures and projects as part of its internal oversight processes. These reviews verify that customer utilities comply with the Implementation Manual, each utility's Energy Conservation Agreement, and

specifications in BPA's reporting system. As such, some of the work BPA already does as part of its oversight processes is similar to certain components of the Evergreen team's evaluation (e.g., file reviews for sampled projects). The evaluation team will continue to work with BPA COTRs and the engineering staff that support the oversight process as much as possible to coordinate efforts and communicate to sampled customer utilities. This includes meeting periodically to share focus areas, processes, and results with the goal of minimizing impacts on customers and maximize information sharing between COTRs and evaluators. For specific evaluation projects, when evaluation samples are selected, the evaluation team will cross-reference each project with available project information collected by the COTRs to minimize effort for customer utilities.

2.4.3 UTILITY CUSTOMER AND END USER CONTACT PROTOCOL

The Evergreen team will adhere to the detailed end user and utility contact protocols provided in the appendices. These protocols describe how the evaluation team must contact customer utilities and end-users across all data collection efforts. For each project, the Evergreen team, in coordination with BPA, will create a project-specific timeline, and a communication protocol document will be developed.

Generally, where utilities must provide data to BPA or where end-user customer contact is required, the following communication principles are used:

- Utilities are notified of their projects included in the evaluation prior to the start of evaluation activities and provided with clear information on samples, timelines, and requirements.
- BPA provides opportunities for utilities to understand the details of the evaluation plan and data request.
- BPA gives utilities a reasonable timeline to collect project and billing data, and uses escalation protocols if deadlines are missed, which include the BPA COTR and Account Executive.
- BPA provides at least four weeks of notice to utilities prior to any end-user contact, including phone surveys and site visits.

3 APPENDIX A: CUSTOM AND C/I/AG LIGHTING EVALUATION

The evaluation will represent the population of measures installed during FY2020 and 2021[12] for the Custom and C/I/Ag Lighting portfolio (including non-residential lighting and custom[13] measures). There are three evaluation objectives:

1. Estimate first-year kWh savings for the Custom and C/I/Ag Lighting portfolio and for separate portions of the portfolio as needed to understand the savings performance of important program delivery channels.
2. Develop recommendations on M&V procedures, including when savings can be reliably estimated, for custom measures, using the BPA M&V Protocol Selection Guide, including the protocol called *Engineering Calculations with Verification* (ECwV).
3. Conduct the evaluation as a continuous process throughout the two-year period, establishing a model for continuous evaluation in future years.

SAMPLE DESIGN

The sampling unit of this study is a measure, defined as a unique Technology/Activity/Practice for a single project.[14] Table A-1 shows the number of measures and savings associated with each domain for 2018. The domains are categorized by similar delivery approaches: utility type (Option 1/Option 2), measure type (lighting/non-lighting), and sector (industrial/non-industrial)

For this design, we have assumed that future years will have a similar distribution of measures and savings. When sampling is conducted, we will select a savings stratified random sample from measures that completed invoicing in the previous year. This will allow the evaluation to be conducted on a rolling basis, meeting the third objective for this evaluation.

The sample sizes shown in the table will achieve a sampling precision of +/- 10% at a 90% confidence level for the entire portfolio of custom and C/I/Ag measures paid in each fiscal year. The precision of estimates for each domain will be +/- 20% at an 80% confidence level.

[12] Based on when invoicing process is complete. Completion dates for projects may also include FY2019 due to natural delays from project completion to utility invoicing to BPA.
[13] For Option 2 utilities, both lighting and custom projects are reported to BPA through the custom project pathway. For this evaluation, Option 2 custom projects are technically those projects with non-lighting end uses.
[14] For uniformity of evaluation approach, evaluation and project resource management, and cost control, sampling is based on measure.

Table A-1: Measures, Savings, and Sample by Delivery Channel

			Number of Measures		Savings		Sample Size
Utility Type	Evaluation Category	Sector	Count	%	aMW	%	Total
Option 1	Lighting	All	4,147	71%	13.4	37%	35
Option 2	Lighting	All	1,218	21%	10.6	29%	35
Option 2	Custom	Industrial	27	0%	0.9	2%	18
		Non-Industrial *	224	4%	2.4	7%	25
Option 1	Custom	Industrial	122	2%	6.7	19%	25
		Non-Industrial *	71	1%	1.1	3%	18
	ESRP	Federal	9	0%	1.1	3%	9
		Total	**5,818**	**100%**	**36.1**	**100%**	**165**

* May include residential projects.

The sampling will leverage the 2015 Site Specific evaluation's sampling tool. The sampling will be conducted with a conventional optimum allocation stratified design based on reported savings for the measure. Each domain will be divided into four or five strata, with certainty selection in the largest savings strata. Each stratum will be defined by an upper and lower bound of project savings. The certainty stratum helps control the variance in the estimate of total domain savings and usually will include two or three projects with the largest savings. We will also define an excluded stratum that contains very small savers; generally, this is the group of projects that collectively account for less than 2 percent of the domain savings. A simple random selection will be used up to the optimal sampling fraction in each of the other strata. All utilities and projects will be included in the sample pool (i.e., no utilities or projects will be excluded). Because of the staged evaluation approach, there will be at least a year between evaluation requests of individual utilities.

We have proposed sample sizes that should yield a relative precision of +/- 10% at a 90% confidence level for each domain. The required sample size to achieve this goal depends on the number of projects comprising a domain. Once the domain population is above 500, the number required is relatively constant. As the population gets smaller, we must select a larger fraction of the projects in order to achieve this goal. Sample size is also determined by the variation in savings across the projects, and we will be able to refine these sample sizes once the reported savings are available for each domain.

In addition to this analysis, the Evergreen team calculated that control savings constitute less than 5 percent of the portfolio savings, which is the threshold recommend by the RTF's lighting research strategy. Therefore, this evaluation plan does not include research on controls savings fractions.

TIMING

To enable an evaluation that provides timely results on a consistent, rolling basis, the evaluation will conduct one domain-specific study approximately every six months over a four-year period, sampling one year of prior history. Table A-2 outlines the proposed timing of each domain's study.

Table A-2: Timing of Custom and C/I/Ag Lighting Evaluation Study Domains

Utility Type	*Evaluation Category*	*Sector*	*FY2020*				*FY2021*				*FY2022*				*FY2023*			
			Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Option 1	Custom	Industrial																
Option 2	Custom	Industrial																
Option 1	Lighting	C/I/Ag																
Option 2	Lighting	C/I/Ag																
Option 1	ESRP	Federal																
Option 2	Custom	C/Ag																
Option 1	Custom	C/Ag																

DATA COLLECTION

Our general approach to evaluation data collection will be to fully leverage the data collected by BPA and utility staff throughout the process of developing each project and to only collect additional data if needed to achieve reliable estimates of savings for the sampled measures. We will collect the necessary data as follows:

- **File review.** The file review will involve extracting all project information relevant to savings estimation including measure descriptions, baseline or efficient condition inputs, reported savings values, and the final version of the M&V model.
- **Telephone/email discussion with project engineers.** The project engineers (BPA, utility, or ESI) are another possible source of data. As needed, we will contact them by telephone or email to obtain information needed for the evaluation that was not found in the project files.
- **Telephone/email discussion with end users.** In some cases, it may be necessary to obtain information from the end user via telephone or email contacts.

- **Site visits.** Based on the file review and discussions with internal engineers, we may determine that more information will be needed from inspection of affected systems and equipment, in-person interviews with operation staff, review of electrical and mechanical plans, inspection of control settings, review of manufacturers' specifications, and one-time measurements.
- **Affected system trend metering.** For custom projects, if the metering data are not sufficient, additional metering data will be collected. In some cases, this might be billing or interval premise electric metering data.
- **Weather.** If weather data from the file review are not adequate, actual and/or typical meteorological year (TMY) weather data will be acquired for the most appropriate National Oceanic and Atmospheric Administration (NOAA) weather station.

CUSTOM AND C/I/AG LIGHTING ANALYSIS

CUSTOM MEASURES

We will estimate savings for sampled custom measures as described below.

SELECT RELIABLE EVALUATION MODEL

Our starting point in estimating savings will be a review of the M&V model. The first step will be to determine whether the M&V model conforms to the most appropriate BPA M&V protocol. It is important to note that determining compliance with a BPA M&V protocol is just the first step in reviewing an M&V model. The BPA M&V protocols provide guidance on general approach and specific examples, but they do not provide detailed specifications for every type of efficiency improvement and affected system or equipment. Once the BPA M&V protocol compliance is determined, we will then examine the savings calculations in more detail to determine whether they provide the best practical estimate of savings.

We will conduct the model review during the file review. During this review, we will determine, relying on professional engineering judgment, whether the model, if provided with reliable input data for the savings determinants, will provide sufficiently reliable estimates of savings. An unreliable model would have a high likelihood of greater than 20 percent uncertainty in the overall savings because of misspecification. For example, if a small VFD measure in an industrial plant relied on whole-facility billing analysis to estimate the savings, we might consider this application inappropriate because of its high unreliability. As part of the evaluation, we would specify an evaluation model—such as post-metering for several weeks applied to manufacturers' pump curves—that would be more likely to provide reliable savings. The outcome of each model review will be a decision on whether to use the M&V model or replace it with another model when we estimate savings for the evaluation. This decision will affect what is done in subsequent steps described below.

Other areas germane to the model review include whether or not the model addresses significant measure and/or end use interactions, and whether or not it adequately establishes the proper baseline (current practice or pre-condition as defined in the RTF Current Operative Guidelines[15] and the BPA M&V Protocols). We will consider

[15] https://nwcouncil.app.box.com/v/2018RTFOperativeGuidelines

interactions significant if it is likely that the interactive effect exceeds 10 percent of the measure savings.

If the M&V model is determined to be reliable, then we will adopt it as the evaluation model, and improve its input data if necessary. If the M&V model is found to be unreliable, and thus not suitable to serve as the evaluation model, then we will either enhance or replace it. Enhancement would likely mean adding or replacing certain features, such as measure and end use interactions, while replacement would entail a wholesale change in approach, such as using the Excel-based Energy Charting and Metrics (ECAM) tool instead of a bin model.

The approach above assumes that adequate measure information is available, and that in particular, data and analysis files are transparent and accessible. For example, if an M&V model is only available as a PDF file, then it is generally impossible for the evaluation team to assess the underlying algorithms and formulas for appropriateness and accuracy. In such a situation, it may become necessary to reconstruct the original model or build an alternative one. Consequently, missing or inaccessible M&V models and supporting data can lead to significant costs, not only for the evaluation team to develop models from scratch and reproduce data where feasible, but also in terms of additional data collection burdens on customers. When we encounter instances where the M&V model is not functional, we will work closely with BPA and/or the Option 2 utilities to obtain the necessary information.

ASSESS DETERMINANT RELIABILITY

Once the evaluation model—either the M&V model or a more reliable replacement—has been selected, we then must consider each of the model inputs and determine what level of data collection is needed to support a sufficiently reliable savings estimate. In general terms, as laid out in the RTF guidelines, key determinants of savings include, but are not limited to:

1. Hours of operation
2. Equipment efficiency at full and part-load operation
3. Control sequence and settings
4. Outside air temperature, or other weather parameters
5. Production rate and schedule
6. Building occupancy
7. Time of day

During the file review, we will develop a list of critical determinants for that particular project, where critical is defined as having a significant (possibly 10 percent or more) impact on the calculated savings. We will then find the corresponding values used in the evaluation model, assess the data and/or documentation underlying those values, and determine whether we consider those values reliable. This will involve some engineering judgment. To the extent that sampled measures involve similar systems, equipment, and modeling techniques, we will ensure that consistent judgments are applied.

For instance, we may determine that hours of operation are a critical determinant for a fan control measure at an industrial facility with a weekly schedule. If the evaluation model incorporated pre- and-post metering for two weeks on a random selection of

affected fans, then we may conclude that the determinant value is reliable. However, if the metering only spanned two hours, we may conclude it was unreliable, and therefore, additional onsite metering would be necessary to develop a reliable determinant value. For each measure, we will document our rationale for establishing whether determinants are reliable or not, and how the reliability will be improved if necessary. We will compare across sampled measures to ensure consistency, as well as to identify overarching trends and issues. Preliminary evaluation assessments for measures will be provided to the BPA evaluation team, so they can provide input.

For unreliable critical determinants, we will assess what level of data collection involving the end user would be necessary to obtain reliability for that determinant. In order of cost and complexity, these levels would be (1) telephone/email interview, (2) site visit, and (3) metering. The highest level across all unreliable critical determinants would then determine the level of data collection for the measure.

COLLECT SUPPLEMENTAL DATA

Based on the previous step, we will develop a data collection plan for each sampled measure that will establish how data for each unreliable critical determinant is to be obtained. Our general intent will be to use the least costly and intrusive approach to obtain sufficiently reliable values—starting with telephone interviews, proceeding to a site visit if necessary, and then performing metering in the most critical instances.

It is conceivable that certain measures could require multiple metering rounds. Hypothetical examples include (a) a fruit processing facility with seasonal production schedules, or (b) a complex HVAC controls project that required separate summer and winter data sets to assess cooling and heating performance, respectively. Such instances would likely be rare and would be kept to a minimum because of the inconvenience to the end user, as well as the cost to the evaluation.

The data collection plan would outline for the BPA evaluation team the types of data to be collected prior to and during the site visit when needed. For example, a site visit may involve interviews to find out about production seasons, coupled with collection of nameplate data and short-term metering. It would map out a work sequence to collect data efficiently, with minimal impact on the end user. The plan would also include unit sampling in situations where the measure consists of many pieces of equipment.

RUN EVALUATION MODEL

If the M&V model is deemed appropriate to serve as the evaluation model, and the critical determinant values are deemed reliable, then this step will essentially be a quality control check. If the file review uncovered any clerical or procedural errors that led to a mistaken savings value being reported, then those errors will be corrected, and the proper values recorded for this evaluation.

Otherwise, analysis will consist of running the evaluation model with reliable determinant values obtained through evaluation data collection.

ESTIMATE SAVINGS USING ENGINEERING CALCULATIONS WITH VERIFICATION (ECWV)

We will also estimate savings for each sampled measure using the ECwV protocol. Our lead engineer for the site will create a version of the site data that contains only the data needed for ECwV. In general, this will eliminate trend data obtained from sub-

metering and any conclusions reached by the analysis of such data. It may also be necessary to substitute a different savings estimation model. Our team will use the ECwV model to estimate savings and compare that to the best practical evaluation model results. We will determine the relative reliability of the two estimates.

LIGHTING MEASURES

A large part of the sample will be lighting measures. In most cases, the M&V model will be some version of the BPA lighting calculator or a similar lighting calculator developed by an Option 2 utility. We will base the evaluation model on the most recent version of the BPA lighting calculator. As necessary, the model will be modified in order to accurately represent the baseline condition for specific measures. Efficient case conditions will be determined by documentation provided by utility staff, or if that is not reliable, by a site visit and inspection of the lighting equipment. We will obtain reliable estimates of operating hours by interviewing building operators and occupants or by inspecting settings on time clocks and other control systems. We will begin this work with a comparison of the RTF protocol and the BPA lighting calculator to determine whether baseline and efficient conditions are treated in a similar fashion for various types of fixtures, lamps, and controls. We will report on important differences.

TREATMENT OF INTERACTIVE MEASURES

Savings achieved by one measure can affect the savings of another measure—for example, an HVAC upgrade and improvements to lighting that affect the same spaces within a building. The change in lighting increases the heating load and decreases the cooling load. How much is saved by the HVAC upgrade could be significantly different with and without the lighting change. Thus, the order in which savings are estimated can make a difference. If the two improvements occur as part of separate projects that were completed at different times, this should not be an issue for this evaluation. Whichever measure we sample, we will account for the baseline conditions of the affected systems and equipment. If the HVAC is sampled and the lighting occurred first, our evaluation model will capture the lighting characteristic as part of the baseline conditions.

A problem may arise if one or more projects are completed at essentially the same time. Using information from the reporting system, we will determine whether this occurs for any of the measures in our sample. If it does, we will obtain documentation for all the interactive measures at the end user site so that we can determine how the M&V models accounted for the interactions. We are looking for the measure order that was assumed in estimating each measure's savings. We will use the same measure order in estimating the evaluation savings.

TIME-BASED VALUE OF SAVINGS AND COST EFFECTIVENESS

There are a number of strategies for characterizing the time-based value of savings for the sample of measures. For this evaluation, we will assign load shapes to individual measures, as the current custom project calculator uses load shapes by sector. Using ProCost, we will assign each measure, via its BPA Technology/Activity Practice (TAP) reporting code, to one of the RTF savings shapes. We will then calculate cost effectiveness and peak savings based on the generic calculator and project-specific ProCost analyses and report on any differences.

STUDY AND PORTFOLIO ANALYSIS

Once data collection and analysis are complete for the sample, we will analyze the results and estimate electric savings and cost-effectiveness for each study and for the portfolio. In addition, for custom measures, we will determine when the ECwV protocol provides a reliable estimate of savings.

FIRST-YEAR KWH SAVINGS

We will estimate first-year savings for each delivery channel using the evaluation model results for the sample, weighted to reflect the probabilities of selection. Each sampled measure has a weight that is the inverse of the sampling fraction for the stratum from which it was selected. We will adjust that weight to account for any instances where the number of measures evaluated for a stratum changed from the design. This final weight will be used in forming delivery channel estimates of savings.

RELIABLE SAVINGS FROM ECWV

We will prepare two estimates of savings for each custom measure in the sample. The first estimate will be based on the best practical model using all data available for mode inputs, including trend logs obtained via sub-metering or from customer control systems. The second estimate will be prepared using the ECwV protocol, which in general does not utilize trend logs, and uses a simpler engineering model that is consistent with available data. We will compare the two savings estimates and examine the assumptions made in developing the simpler ECwV model and the inputs to that model. We will determine the relative reliability of the ECwV estimate. We will use these comparisons in developing guidelines for when the ECwV model provides sufficiently reliable savings estimates. We will also compare the program and evaluation estimates when both are done with ECwV to determine whether they are systematic differences and to identify ways for the program to improve the reliability of its ECwV savings estimates.

REPORTING

We will prepare a report on each domain as it is completed. The reports will document the methodology, findings, and recommendations of the domain's evaluation. The report will not contain any information that could be used to identify the end users that participate in the evaluation. Further, it will not contain any utility-specific findings or recommendations.

The report will be consistent with the content, transparency, and comparability guidance found in the RTF's Program Impact Evaluation guidelines. We expect that the report will have the following structure:

1. Executive Summary
    a. Findings
    b. Recommendations
2. Introduction and Background
3. Objectives
4. Methodology
    a. Data Collection
    b. Savings Estimation
    c. Delivery Channel and Portfolio Savings Estimation

5. Findings
    a. First-Year kWh Savings
6. Recommendations
    a. Opportunities to Improve M&V Savings Estimates
7. Technical Appendices and Data Products

APPENDIX B: STRATEGIC ENERGY MANAGEMENT SAVINGS

BPA began offering its Energy Management (EM) Program to industrial facilities in 2010. Through the program, BPA provides long-term energy management consulting services to educate and train industrial energy users to (1) develop and execute a long-term energy-planning strategy and (2) permanently integrate energy management into their business planning. BPA's EM Program was one of the nation's first large-scale deployments of a strategic energy management (SEM) program in the industrial sector, having engaged 74 projects by the end of 2018.

An in-depth evaluation of the EM Program was completed in 2016. However, two issues were not addressed by this evaluation. These two issues are addressed in this additional evaluation of SEM savings from the EM Program.

EVALUATION OBJECTIVES

This study has two objectives:

1. **Investigate persistence of SEM changes.** The first objective of this study is to determine how long SEM-initiated changes are operational. To the extent possible, we want to determine whether changes are abandoned after the participants leave the program.
2. **Verify capital savings deduction.** In many cases, industrial customers participating in the EM Program also install capital efficiency measures under other BPA programs. In the last impact evaluation, SEM savings (from operational and maintenance actions) were estimated by deducting the capital measure savings from the total facility savings. Thus, SEM savings were the residual. The evaluation did not verify the program savings claims for those capital measures, and there is value in evaluating these measures for reliability. There may have been errors and incompatible savings estimation procedures that caused the capital savings to be incorrect, and thus the SEM residual to be incorrect. Examples of incompatible savings estimation procedures include the use of current practice baseline assumptions in the BPA lighting calculator and the use of typical weather conditions instead of actual weather in estimating HVAC savings. The second objective of this study will be to verify capital measure savings and determine whether errors or incompatibilities in capital measure savings create significant errors in SEM savings. We will apply procedures from Appendix A in verifying the savings from each of the capital measures. Similar to above, we will also assign best-available load shapes and analyze the savings with ProCost for capacity and cost-effectiveness.

SAMPLE DESIGN

We will separately select two samples of EM participants for analysis. Each will serve one of the two objectives for this study (capital savings deduction or persistence of SEM changes). However, the population of EM participants is small, and it is likely that the two samples will overlap, so they need to be closely coordinated to avoid undue burden on the sampled participants.

For the first objective (persistence of SEM changes), we will randomly sample (~15) EM participants from those that were active in the program in 2015. Some of these may no longer be active. We will investigate the persistence of changes caused by the program

within the engagement period. For those who are no longer active, we can investigate persistence of changes beyond the engagement period.

For the second objective (capital saving deduction), we will randomly sample (~10) EM participants that also completed capital projects. We will work with Energy Smart Industrial (ESI) program staff to identify all the EM participants that also completed capital projects. We will select a simple random sample from this group. We do not expect to exclude any utilities from the potential sample and expect to focus on capital projects completed in the past two years.

DATA COLLECTION

Our data collection activities will be tailored to meet the needs of the two evaluation objectives. Some EM participants will be in both samples. These will be identified so that we can tailor the data collection activities to minimize the burden on these participants.

The data collection activities for the first objective (persistence of SEM changes) will be as follows:

1. **Change inventory.** For each sample participant, we will work with program staff to create a list of significant SEM-initiated changes and the date that each became operational. We will focus on the changes that the program staff believe account for a large portion of the SEM savings. Ideally, we will be able to focus on 20 percent of the changes that account for 80 or more percent of savings. Reducing the number of changes, we investigate for each participant will reduce the amount of their staff time needed for this research.
2. **Change status measurement planning.** We will review the change inventory and determine the least intrusive, but reliable, method for determining the status of each change and if the change is no longer operational, when it was abandoned. It may be possible to make this determination via telephone calls with participant staff or their vendors. If not, we will plan to observe evidence of the change during a site visit. For changes that are no longer operating, we will have to rely on participant staff or their vendors to tell us when the change was abandoned. For changes that are still operational, we will ask the participant staff to estimate how long they believe it will remain operational and whether they actively monitor its status (if that is appropriate).
3. **Collect change data.** With the assistance of ESI program staff (and utility staff as relevant), we will contact the sampled customers and collect the needed data regarding the status of each change and the dates that any changes were abandoned.

The data collection activities for the second objective (capital savings deduction) will be as follows:

1. **Capital project data.** We will collect all program documents and other project data as is described under the data collection section of Appendix A. These data will be used to develop engineering estimates of capital project savings. In addition, we will work with the program staff to determine the operational date for each project, if this is not clear from the project files. The operational date for each project will be critical to the regression estimates of savings described under Site-Specific Analysis.

2. **MT&R data and models for each site.** We will obtain these data and models from program staff. These data and models will be used in attempting to estimate capital measure savings using regression techniques.

SITE-SPECIFIC ANALYSIS

Our site-specific analysis activities will be tailored to meet the needs of the two evaluation objectives.

The activities for the first objective (persistence of SEM changes) will be as follows:

1. **Classify changes.** We will examine the list of changes and develop a scheme for classifying these changes. The classification may be based on factors such as type of affected equipment, whether the change affected maintenance or operational practices, whether the change was associated with support services (HVAC, lighting) versus production processes, or whether the change involved installation of any new devices such as sensors, actuators, or controls.
2. **Estimate change lifetimes.** We compute the time between initiation of the change and its last operational date. Some changes may still be operational as of our survey. For these, we will rely on the participant's opinion about its likely lifetime. We may adjust these participant responses, based on whether the participant is actively monitoring the status of a change.
3. **Summarize lifetimes.** We will summarize change lifetimes into one or more categories. To the extent possible, we will examine whether these lifetimes are reduced after the program engagement ends. We will look for evidence that customers continue monitoring the status of changes after the engagement. We will also look for differences in the customers' expected lifetimes for changes that are still operational, comparing customers that are still engaged with the program versus those that are not. Please note that this information would inform a policy decision by BPA on programmatic lifetime. That is, evaluation will estimate lifetimes for sampled measures and projects; from there, BPA would likely use it to inform an average program lifetime.

The activities for the second objective (capital savings deduction) will be as follows:

1. **Engineering estimates of capital measure savings.** We will follow the procedures described in Appendix A in estimating savings for each capital measure. This may result in a new model for some sites if that is necessary to achieve a reliable savings estimate. Inputs to the final selected engineering model will come from the best available sources. We will then adjust the inputs to reflect existing conditions baselines and actual weather. This version will estimate the change in metered use that would be caused by the measure and will be compatible with the total facility savings estimated by the Monitoring, Targeting and Reporting (MT&R) model.
2. **Compute SEM savings.** We will have three estimates of savings for each capital measure: one provided by the program, one based on our final evaluation estimate as described in Appendix A, and one based on the evaluation estimate but adjusted for existing conditions baseline and actual weather. We will deduct each of these from the MT&R estimate of savings to create three estimates of SEM savings. We will analyze the differences between these SEM savings estimates.

3. **Regression estimates of capital measure savings.** We will enhance the MT&R model with information about the timing of capital measures and attempt to estimate the capital measures savings. We will compare these estimates to the engineering estimates and determine when this approach would be practical. When it is practical, it would be a less expensive method for estimating SEM savings.

REPORTING

We will develop two reports, one focused on each of the two study objectives. Our first report will document the methodology and findings related to the first objective, persistence of SEM changes. We will provide recommendations on how the program can estimate the lifetime of SEM savings based on the types of changes made by each participant. For the second report, we will describe the methodology and findings related to capital measure savings estimation. This report will provide recommendations on how the program can improve its estimate of the capital measure savings before it estimates SEM savings and will align with the custom project evaluation described above, wherever possible. The procedure for correcting these estimates may be based on regression analysis, if we find that that is practical for certain MT&R models.

APPENDIX C: COMMERCIAL BPAQ HVAC

BPA offers commercial BPA Qualified (BPAQ) HVAC measures for ductless heat pumps and air-source heat pumps (both conversions and upgrades). In FY2020-21, these measures will be evaluated using a billing regression model. The billing regression provides an estimate of realized savings that takes into account all the factors that may be contributing (or detracting) from savings. In this respect, it has the advantage of estimating savings based on real-world operating conditions. By its nature, the billing regression utilizes the existing pre-participation conditions as the baseline instead of assigning a standard practice baseline.

SAVINGS/POPULATION

The table below summarizes the number of utilities and projects for 2018 based on BPA's reporting data. Participation is limited, with most of the sites installing ductless heat pumps (DHPs).

Table C-1: 2018 BPAQ Commercial HVAC Participation

	Air-Source Heat Pumps; Conversion	*Air-Source Heat Pumps; Upgrades*	*Ductless Heat Pumps*	*# of Utilities*
2017	2	13	95	38
2018	10	15	124	43
Total	**12**	**28**	**219**	**57**

In the summer of 2019, the Evergreen team requested project documentation and billing data for 2017 and 2018 projects. The team received data for 92 DHP and 13 air-source heat pump projects. The evaluation team will request the 2019 form data for BPAQ sites at the beginning of the evaluation to bolster the sample available for the billing regression.

SAMPLE SIZE

For this evaluation, we will utilize the full population of participants from 2017-2019 in the billing analysis. As discussed above, the current data indicate that there are approximately 105 measures available for 2017 and 2018 and an unknown quantity for 2019. The analysis, and the number available for the analysis, will decrease depending on how many customers we are able to obtain billing data for over an adequate pre-participant and post-participation period. Ideally, we will have a minimum of 12 months of pre-installation and post-installation data. Note that we will still utilize the 2019 sites as they will help bolster our sample and provide additional sample points in the pre-participation period, even if they do not have the desired months of post-participation billing data.

APPROACH

Our approach for these projects is to conduct several variations on the billing regression model, as discussed below. We will likely need to focus on the ductless heat pump measures given that they have the largest number of sample points; however, we will

attempt to include the other heat pump measures either in the overall model (with variables to flag these sites) or else as a separate model.

Our first modeling effort will involve a relatively simple analysis of heating and cooling loads based on billing data. This will involve some simple correlation calculations in addition to more standard billing regression methods that will attempt to identify the share of load devoted to either heating or cooling as a function of outside temperatures.

We also propose to conduct a pooled billing regression model using a fixed effects model specification similar to the work we have already done for the PTCS measures. This model has some advantages in that each customer is assigned a unique constant term that helps control for site-specific characteristics. The pooled data provide a greater analysis sample, which will be important given the smaller number of participants available for this program. We will use a basic fixed effects specification as a starting point:

$$kWh_{it} = \alpha_i + \beta_1 Post_{it} + \beta_2 Part_i + \beta_3 Part_i \times Post_{it} + \beta_4 CDD_{it} + \beta_5 HDD_{it} + \beta_6 Post_{it} \times CDD_{it} + \beta_7 Post_{it} \times HDD_{it} + \varepsilon_{it}$$

$Where$:

$kWh_{it} =$ Electricity usage by the i^{th} home in the t^{th} time period

$Post_{it} =$ Indicator variable for month in the post − participation period

$Part_{it} =$ Indicator variable for BPAQ HVAC installed in time t

CDD_{it}= Cooling degree days

$HDD_{it} =$ Heating degree days

$\alpha, \beta =$ Coefficients to be estimated in the model

$\varepsilon_{it} =$ Random error term

Variations of this model will also be explored, including interacting the participant variables with the weather variables and including separate indicator variables for each month. We will also explore comparing this initial fixed-effects model with the results of a site-specific regression model, using either ECAM or similar analysis methods.

We will also do model variations including the 'post only' model that has been utilized in other BPA impact evaluations:

$$kWh_{i,t} = \beta_1 Part_i + \beta_2 CDD_{i,t} + \beta_3 HDD_{i,t} + \sum_{k=1}^{12} \beta_k Month_k + \sum_{j=1}^{12} \beta_j Month_j \times kWh_{t-12} + \varepsilon_{i,t}$$

Where:

$kWh_{i,t} = Average\ daily\ kWh\ of\ customer\ i\ in\ month\ t$ of the post period

$Part_i = Indicator\ variable\ that\ equals\ 1\ if\ customer\ i\ is\ a\ participant, else\ 0$

$CDD_{i,t} = Average\ cooling\ degree\ days\ for\ participant\ i\ in\ month\ t$

$HDD_{i,t} = Average\ Heating\ degree\ days\ for\ participant\ i\ in\ month\ t$

$Month_k = Month_j = Indicator\ variable\ that\ equals\ 1\ for\ month\ k\ or\ j, else\ 0$

$kWh_{i,t-12} = Average\ daily\ kWh\ of\ customer\ i\ in\ month\ t\ of\ previous\ year$

$\beta_1, \beta_2, \ldots = Parameters\ to\ be\ estimated\ in\ the\ model$

$\varepsilon_{i,t} = Random\ error\ term$

APPENDIX D: CLARK PUBLIC UTILITIES THERMOSTAT PILOT

Clark Public Utilities has begun a smart thermostat pilot program in 2019 that targets multi-family customers. The thermostats are installed in the individual units, focusing on low-income residents as designated by the Vancouver Housing Authority. The program allows for the replacement of up to five thermostats per unit based on the customer's preference and the most heavily used areas. The type of thermostat installed through the pilot is the Honeywell #L7235A1003/U.

The impact evaluation will focus on determining realized savings from the thermostat installs by examining the energy consumption before and after the thermostat installations while controlling for weather conditions. Depending on the results of the initial billing regression, we will discuss with BPA the potential benefits and feasibility of conducting a tenant survey to determine other factors that may affect energy consumption (e.g., change in occupancy, changes in how thermostats are used).

SAVINGS/POPULATION

From Clark Public Utilities, the total current pilot participation is as follows:

- 1,800 thermostats installed (707 multi-family units)
- 1,239 units scheduled, plan to finish by March 2020
- 2.5 thermostats per unit on average

The evaluation will plan to begin approximately one year after the majority of sites have been completed. Therefore, delays in the implementation schedule may lead to evaluation delays.

SAMPLE SIZE

Based on the amount of expected participation, we will be utilizing the entire participant population to estimate savings, rather than drawing a sample. Using all the available data will allow us the maximum flexibility for the analysis. It will also eliminate the need to return to Clark Public Utilities for a second data request if we ask for all available data at the beginning of the evaluation.

APPROACH

For the thermostat pilot, we will use a similar billing regression approach as discussed above for the BPAQ HVAC measure assessment. For the thermostat pilot, the initial model will be a fixed effects billing regression that will allow a customer-specific constant term to be included to control for factors that are specific to individual customers. We will also explore a variation where constant terms are also included for specific multi-family complexes that have multiple tenant participants. Variables will also be included to account for the number of thermostats installed in each unit. One possible specification is:

$$kWh_{it} = \alpha_i + \gamma_k + \beta_1 Tstat_{it} + \beta_2 Post_{it} + \beta_3 Part_i + \beta_4 Part_i \times Post_{it} + \beta_5 CDD_{it} + \beta_6 HDD_{it} + \beta_7 Post_{it} \times CDD_{it} + \beta_8 Post_{it} \times HDD_{it} + \varepsilon_{it}$$

Where:

$kWh_{it} =$ Electricity usage by the i^{th} home in the t^{th} time period

$\alpha_i =$ Individual constant term for home i

γ_k = Individual constant term for multifamily complex k

$Tstat_{it}$ = Number of smart thermostats installed by the Pilot at home i

$Post_{it}$ = Indicator variable for month in the post − participation period

$Part_i$ = Indicator variable for Pilot participants

CDD_{it}= Cooling degree days

HDD_{it} = Heating degree days

β = Coefficients to be estimated in the model

ε_{it} = Random error term

Additional variations on this model will be explored based on available data and reviewing standard goodness of fit statistics. If AMI data are available, we will also develop separate billing regression models that estimate savings based on average daily load shapes.

Currently, the participation data are available on paper applications that are then converted electronically by Clark Public Utilities program staff. We will request these data along with 12 months of pre-participation billing data. We will also request 12 months of post-participation billing data, with the understanding that there may not yet be 12 months in the post-installation period at the time the evaluation commences. At a minimum, we would like the post-installation period to include as much of the winter and summer months as possible and will adjust the timing of the evaluation accordingly. Utilizing the entire participant population for the analysis (rather than just a sample) may help alleviate some of the issues around having a limited number of post-installation months if there is enough variation across customers.

APPENDIX E: EVALUATION PLANNING FOR UES MEASURES

Table E-1: Residential UES Measures

Technology/Activity/Practice	*RTF Status*	*aMW*	*Evaluation Pros*	*Evaluation Cons*	*Decision*
Showerheads	RTF Approved-Planning	0.3	Regional need – no known research	WA legislation	Watch: WA legislation + future size
Lamps	RTF Approved-Proven	5.5	Big (2018)	Large savings drop, good evaluation results in future	Watch: future savings
Fixtures	BPA Qualified-Structural	0.5	Will continue, BPAQ structural	Portfolio shifting away	Watch: future savings
Heat Pump Water Heaters	BPA Qualified-Structural/RTF Planning	0.3	Regional need – no known research	Looks like Momentum Savings	Exclude
Power Strips	RTF Approved-Planning	0.2	Regional need – no known research; ETO research completed then stopped Tier 2	Not strategic importance to portfolio. BPA only offers Tier 1 (new)	Watch: future savings
Thermostatic Shut-off Valves	RTF Approved-Planning	0.1	RTF Planning		Watch: Showerheads
Windows	BPA Qualified-Structural/RTF Proven/RTF Planning	0.5	Opportunity: Low-E storm windows BPAQ structural and RTF Planning	Billing analysis several years ago, Low E savings small 2018	Watch: Low E future savings

Technology/Activity/Practice	*RTF Status*	*aMW*	*Evaluation Pros*	*Evaluation Cons*	*Decision*
Insulation	BPA Qualified-Structural/RTF Proven/RTF Planning	0.3	BPAQ structural and RTF Planning	Planning for HP baseline, billing analysis several years ago	Exclude
NEEM Certified Home	RTF Approved-Planning	0.1	RTF Planning	NEEA conducting research	Exclude
Low Income Residential Weatherization	RTF Standard Protocol	0.1		Recent evaluation work	Exclude
Thermostats	RTF Approved-Planning	0.1	RTF Planning	NEEA conducting research	Exclude
Clothes Washers	RTF Approved-Planning	0.1	RTF Planning		Exclude

Table E-2: Commercial UES Measures

Technology/ Activity/ Practice	*RTF Approval Status*	*aMW*	*Evaluation Pros*	*Evaluation Cons*	*Decision*
Thermostats	RTF Approved-Planning	0.12	RTF Planning regional research opportunity; No regional research being undertaken	Small	Watch: future savings, new thermostat measure
HVAC Control Improvements (VFD)	BPA Qualified/RTF Planning	0.10			Watch: future savings
Power Strips	RTF Approved-Planning	0.00			Watch: future savings
Ductless Heat Pumps	BPA Qualified	0.13	BPA Qualified, opportunity to analyze and support update	Engineers working on this	Include in 2020-2021 plan
Air-Source Heat Pumps; Conversion	BPA Qualified	0.05			
Air-Source Heat Pumps; Upgrades	BPA Qualified	0.02			
Windows	BPA Qualified	0.02			
VRF	BPA Qualified	0.01			
Insulation	BPA Qualified	0.01			
Strip Curtains	RTF Approved-small saver	0.10	RTF non-proven	Importance of grocery in future	Exclude
Motors	RTF Approved-small saver	0.02			

Technology/ Activity/ Practice	*RTF Approval Status*	*aMW*	*Evaluation Pros*	*Evaluation Cons*	*Decision*
Electric Combination Ovens	RTF Approved-Planning	0.01			
Refrigeration Control Improvements (non-VFD)	RTF Approved-small saver	0.01			
Electric Convection Ovens	RTF Approved-small saver	0.00			
Hot Food Holding Cabinets	RTF Approved-Planning	0.00			
Interactive Process Loads System Improvements	BPA Qualified	0.01	Block heaters, DC kitchen ventilation	Very small	Exclude
Clothes Washers	RTF Approved-Planning	0.00		Very small	Exclude

Table E-3: Industrial UES Measures

Technology/Activity/ Practice	*RTF Status*	*aMW*	*Evaluation Pros*	*Evaluation Cons*	*Decision*
Motors/Drives Control Improvements (VFD)	RTF Approved-small saver	0.08		Small	Exclude
BPA Green Motors	RTF Approved-small saver	0.01		Delivery verification recently	Exclude
Interactive Process Loads System Improvements	RTF Standard Protocol	0.00	Block heaters	Very small	Exclude

Table E-4: Agriculture UES Measures

Technology/ Activity/ Practice	*RTF Approval Status*	*aMW*	*Evaluation Opportunities*	*Evaluation Considerations*	*Initial Eval Priority*	*Approaches*
Irrigation						
Nozzle Replacement	RTF Approved-Planning	0.3	Never been evaluated, high priority by engineering (who are focused on new measures), RTF Planning. Could coordinate with Idaho Power.	Small savings, likely difficult to evaluate due to decentralized nature (small measures, large geography, small % of savings)	Medium	Watch: future savings
Regulator Replacement	RTF Approved-Planning	0.2			Medium	
Sprinkler Replacements	RTF Approved-Planning	0.1			Medium	
Drop Tube/Hose Extension	RTF Approved-Planning	0.0			Low	
Goose Necks	RTF Approved-Planning	0.0			Low	
Multi-Trajectory Sprays	RTF Approved-Planning	0.0			Low	
Pump Efficiency Upgrade	BPA Qualified	0.0	Future opportunity?	Very small	Medium	Watch: future savings
Utility Distribution System						
De-Energization	RTF Standard Protocol	0.2	Beyond DV?	Recently evaluated	Low	Exclude
Motors/Drives						
Motors/Drives Control Improvements (VFD)	BPA Qualified	0.1	Future opportunity		Medium	Exclude

APPENDIX F: UTILITY CUSTOMER CONTACT PROTOCOLS

CONTACT PROTOCOLS

The evaluation team will adhere to the following general end user and utility contact protocols for each evaluation that includes approaches that require the team to contact end users and utility representatives. For each project, the Evergreen team, in coordination with BPA, will create a project-specific timeline, and a communication protocol document will be developed. Generally, where utilities must provide data to BPA or where end user customer contact is required, the following communication principles are used:

- Utilities are notified of their projects included in the evaluation prior to the start of evaluation activities and provided with clear information on samples, timelines, and requirements.
- BPA provides opportunities for utilities to understand the details of the evaluation plan and data request.
- BPA gives utilities a reasonable timeline to collect project and billing data, and uses escalation protocols if deadlines are missed, which include the BPA COTR and Account Executive.
- BPA provides at least four weeks of notice to utilities prior to any end user contact, including phone surveys and site visits.

1. UTILITY PRE-NOTIFICATION AND OVERVIEW BROWN BAG

Once the evaluation plan and sample are final (or nearly final), BPA will notify utilities via email that at least one project in their territory has been (or may be) selected in the evaluation sample. This initial email will request the primary utility contact for the evaluation and provide information summarizing the projects, measure groups, and/or approaches for which the utility may be sampled. This email may also include an invite to an overview brown bag and an option to review the evaluation plan.

2. UTILITY NOTIFICATION OF SAMPLE AND DETAILED BROWN BAGS

Once the evaluation plan and sample are final, the evaluation team will provide detailed information to each utility about their sampled sites (e.g., address, completion date, number of units, invoice number) through a secured file transfer protocol (FTP) and provide detailed information on what information is needed, as well as any data templates to be completed. If needed, BPA will organize a kickoff meeting to provide detailed information about the evaluation, its general process, and the contact protocols.

Any utility submitting data directly to the evaluation team may negotiate and execute with the evaluation team a non-disclosure agreement that meets the utility's requirements for protecting end user information.[16] BPA's contract with the Contractor protects data under the language of BPA's existing contract with the evaluation firm.

[16] BPA has a contract with the evaluation firm that requires data protection of the data. Therefore, this NDA may be most useful to utilities that provide data directly to the evaluation team.

3. PROJECT DOCUMENTATION REQUESTS

BPA (and its contractors such as Energy Smart Industrial) will provide the contractor all relevant project information it has, including custom project completion files, lighting calculator files, and COTR oversight documents. If BPA cannot provide the project documentation for samples projects, the utility will be contacted by the evaluation team and the needed files will be included on the sample list. While the focus will be on the required documentation, utilities may provide whatever additional data they collect to the evaluation team.

The evaluation team will provide a timeline for file delivery. The utility (or BPA, if requested by the utility) will upload required files to a secure website. The evaluation team will work with utilities individually to support their data request as much as feasible, including providing support staff to collect (scan and upload) paper files, etc. An extended delivery date may be requested and will be accommodated, if possible.

In order to strive to provide timely and actionable evaluation results, the team has created an escalation protocol to be initiated should data collection efforts become significantly delayed and pose an impact to the schedule. The protocol is:

1. Initial sample emails sent by the evaluation EER with copy to the utility EER and utility COTR.
2. If a utility requests more time, within the agreed-upon time limit, the utility EER and utility COTR are notified.
3. If a utility misses the deadline, then the evaluation EER, utility EER, COTR, and AE are notified of the missed deadline. The utility EER and the utility AE will discuss an approach to the data collection, including potential escalation to utility management.

Additionally, BPA may interview utility project contacts (e.g., project engineer) to understand more about the project, on an as needed basis.

4. BILLING DATA REQUESTS

Billing data refers to energy consumption data by customer and premise for relevant participants. Depending on the measure being evaluated, the template may also include additional data fields to fill out on an “if available” basis, such as for an existing primary heating system.

These data will be collected using a data template Excel workbook. This workbook will include instructions, an example, the data template to fill out, and contact information for any questions that arise.

The evaluation team will provide a timeline for file delivery, which will provide a minimum of six weeks. The utility (or BPA if requested by the utility) will upload required files to the secure website. The evaluation team will work with utilities individually to support their data request as much as feasible, including providing support staff to collect (scan and upload) paper files, etc. An extended delivery date may be requested and will be accommodated, if possible.

Following an initial analysis of the billing data, the evaluation team may request additional data for a select number of sites where the evaluation team finds unexpected results. The evaluation team will work with utilities to facilitate the data transfer.

5. PHONE SURVEYS OR SITE VISITS OF END USERS

If phone surveys or site visits are utilized, utilities will be notified at least two weeks prior to any end user contact. Utilities will be provided with the survey instrument or description of information to be collected from the site. BPA will provide materials to support any advance contact they would like to make with end users, such as advance letters, email, or phone call script. Sending letters to primary contacts prior to a recruitment call has been found to increase the success of end user recruitment. If utilities show interest in sending advance letters, BPA will provide an example letter. BPA will also provide a set of potential **frequently asked questions** to minimize any potential concerns by the end users.