BONNEVILLE POWER ADMINISTRATION

2020-2021 Custom Industrial Impact Evaluation for Option 2 Utilities

1 EXECUTIVE SUMMARY

This document presents the results of an impact evaluation of the Bonneville Power Administration's (BPA's) custom portfolio for industrial customers of Option 2 utilities. This report addresses Domain 2 of a rolling plan with additional domains that will address the entire custom measure and commercial, industrial, and agriculture (C/I/Ag) lighting portfolios (which include all nonresidential lighting and custom measures).

The evaluation objectives were to:

1. **Estimate first-year kWh savings** for the custom industrial portfolio for Option 2 utilities, and for separate portions of the portfolio as needed to understand the savings performance of important program delivery channels.
2. **Develop recommendations on measurement and verification (M&V) procedures**, including when savings can be reliably estimated, for custom measures, using the BPA *Measurement and Verification Protocol Selection Guide and Example M&V Plan* and applying the protocol called *Engineering Calculations with Verification* (ECwV).

1.1 METHODOLOGY

This evaluation represents the population of measures with completion dates between Q2 2020 and Q1 2021, with a small supplemental sample of measures going back to 2019 Q4.[1] The sample design targeted (and achieved better than) a 90/10 confidence level and precision and was developed based on BPA tracking data.

The sampling was conducted with a conventional optimum allocation stratified design based on end use category and reported kWh savings for each measure. The sample size was 22 measures.

The data collection approach included a combination of sources such as file review, site visits, and trend metering. Custom measure analysis was conducted using a multistep process starting with review of the M&V model, collecting supplemental data where needed, running the evaluation model, and estimating savings.

The evaluation team also collected information on the impact of COVID-19 on first-year savings.

Once data collection and analysis were completed for the sample, site-level results were compiled to estimate the electric savings and cost-effectiveness for the domain and individual end uses within the domain using a ratio analysis.

1.2 SUMMARY OF FINDINGS AND RECOMMENDATIONS

The overall results for this domain showed evaluated savings coming in slightly above reported savings, for an overall realization rate (ratio of evaluation savings to reported savings) of 102

[1] The sample design initially specified a 12-month period, from 2020 Q2 through 2021 Q1 and invoice approval dates between June 3, 2020, and June 2, 2021. After the utilities were notified of the customers in the sample, the team identified five refrigeration measures from one of the utilities that had been installed at commercial facilities and then mistakenly assigned industrial measure codes. The sites were removed from the sample frame and replaced with a supplemental sample frame with measures between 2019 Q4 and 2020 Q1. For the end use strata, the evaluation cycle still spans 12 months. However, due to this issue, the timeline for the supplemental sample differs from the rest.

percent, as shown in Table 1. These results are shown for the entire domain, where sample-level results have been 2 percent for a 90 percent two-tailed distribution, which was better than the target design of 10 percent.[2] The overall realization rate of 102 percent is very close to the previous evaluation of this domain,[3] which had a realization rate of 107 percent. applied to the population of projects. The overall sampling precision was

Table 1: Evaluated first-year savings by end use

End Use	Realization Rate	Sampling Relative Precision (90% two-tailed)	Reported Savings (kWh)	Evaluation Savings	
				kWh	Percent of Domain
Compressed Air	99%	5%	4,217,596	4,190,554	46%
HVAC	112%	0%	644,183	723,972	8%
Motors/Drives	103%	33%	427,781	439,983	5%
Process Loads	116%	0%	2,812,583	3,275,749	36%
Refrigeration	38%	11%	447,530	170,813	2%
Water Heating	77%	0%	326,635	251,755	3%
Total	**102%**	**2%**	**8,876,308**	**9,052,828**	**100%**

[2] The realized sample precision was better than the original estimate because there was less variation between evaluated and reported savings than expected.

[3] See results for the Option 2 Non-Lighting Industrial domain from the prior evaluation completed in 2017: https://www.bpa.gov/-/media/Aep/energy-efficiency/evaluation-projects-studies/impact-evaluation-site-specific-portfolio-final-report.pdf.

As shown in the table, the realization rates ranged widely by end use from 38 percent for refrigeration measures to 116 percent for process loads. More than half of projects had realization rates either below 80% or above 120%. On an individual project basis, reliability is a factor for utilities and end-use customers to ensure they are achieving savings and paying incentives as intended. Therefore, there is room to improve the reliability of individual project savings. The recommendations in this report offer specific suggestions that, if taken, should lead to closer alignment of reported savings with evaluated savings for individual projects.

Evaluators observed that BPA *Implementation Manual* and measurement and verification protocols were generally followed correctly and that the difference in realization rates did not result from deviations in procedure. The primary reasons why projects had realization rates not equal to 1 were:

- Savings were not calculated based on operation during the first year of implementation.
- More first year operations data were available during evaluation.
- There were errors in the savings calculation approach.

These observations would not notably impact the overall realization rates for the sampled population but would improve the reliability of individual models.

The evaluation also included an assessment of the BPA *Engineering Calculations with Verification* (ECwV) protocol. The evaluators found that small and medium-sized projects showed little bias in evaluated results using the protocol, and the results using the protocol were similar to evaluated results overall.

The evaluators also collected information on each sampled site regarding any adjustments to facility operation due to the COVID-19 pandemic (e.g., changes in production schedules). Evaluators calculated a second set of savings estimates that were intended to represent what would have happened had the COVID-19 pandemic not occurred. Those results were very similar to the evaluated savings with no adjustments for COVID-19.

The evaluation team developed recommendations resulting from the key findings of this study:

- **Recommendation 1:** BPA should clarify requirements for the basis year of savings in the BPA *Implementation Manual*.
- **Recommendation 2:** BPA should offer training and access for Option 2 utilities to BPA's solutions for common measures.
- **Recommendation 3:** BPA should consider applying ECwV to a wider size range of custom industrial projects.
- **Recommendation 4:** BPA should conduct outreach to Option 2 utilities to promote the use of ECwV as a vetted M&V approach for smaller projects that still fully complies with BPA M&V guidelines.

2 INTRODUCTION

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 utilize a constant savings value for each measure application;
- Custom measures require calculation of savings for each project; and
- Calculator measures have standardized savings estimation algorithms and project-specific parameter values (typically lighting).

The subject of this report is an impact evaluation of BPA's custom portfolio for industrial customers of Option 2 utilities.[4]

2.1 KEY TERMS

See Appendix A for definitions of key terms such as reported savings, measure and realization rate, which are used throughout this report.

2.2 BACKGROUND

Over time, BPA and the Regional Technical Forum (RTF) have developed a series of policies and procedures to provide guidance on how BPA should estimate savings from the projects that comprise its energy efficiency portfolio.

The RTF provides guidance on delivery verification for UES and Standard Protocols, and specifies that 90 percent of the portfolio of savings should be evaluated every four years, with additional guidelines by measure type.[5] Consistent with this, BPA has established policies for impact evaluation that state that within a four-year cycle, impact evaluations should sample from at least 90 percent of savings within the BPA portfolio. [6] Additionally, for custom projects, the policies state that BPA should sample from 99 percent of the portfolio over four years and include sampled projects within a domain that sum up to at least 20 percent of the custom savings.

---

[4] Utilities are categorized as Option 1 or 2 for M&V purposes. For Option 1 utilities, BPA is often involved throughout the project lifecycle by providing technical support for project development, implementation, approval, and M&V. Option 2 utilities provide their own technical support including M&V and custom project quality control, e.g., project proposal and completion report review.

[5] Regional Technical Forum. 2020. *Regional Technical Forum Operative Guidelines for the Assessment of Energy Efficiency Measures*: https://nwcouncil.app.box.com/v/2020RTFGuidelines (see Section 5.2.1).

[6] BPA has developed policies for Standards; Planning; Oversight; Impact and Process Evaluation; and Savings for Custom Projects, Calculators, and UES measures.

In addition, BPA has developed an implementation manual that specifies reporting requirements for energy efficiency programs, and measurement and verification (M&V) protocols specifically for custom measures. [7,8]

The Evergreen Economics team (which includes SBW Consulting and Apex Analytics) conducted a detailed assessment of BPA's portfolio in late 2019 to understand evaluation coverage, priorities, and opportunities. Using the results of the data assessment and guided by the BPA evaluation policies described above, the Evergreen team identified that the highest priority was 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 were most recently evaluated for fiscal year (FY) 2012-2013 savings.

The team developed a rolling plan with several domains to evaluate the entire custom measure and C/I/Ag lighting portfolio (which includes all nonresidential lighting and custom measures). [9,10,11] The evaluation was designed to be an ongoing process, establishing a model for consistently timed evaluation in future years and providing feedback to BPA on the quality of data collection and use of M&V protocols. BPA recently completed an update to the evaluation plan that was based on a review of BPA's FY 2022 energy efficiency program accomplishments that confirmed that the custom and C/I/Ag lighting projects remain the highest impact evaluation priority.[12]

---

[7] Bonneville Power Administration. 2019. *2020-2021 Implementation Manual*. https://www.bpa.gov/-/media/Aep/energy-efficiency/measurement-verification/12-bpa-ci-sem-mv-ref-guide-v10.pdf

[8] 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*. The protocols are currently on their ninth revision and will soon be referred to as the *Commercial & Industrial SEM M&V Reference Guide*, found at https://www.bpa.gov/energy-and-services/efficiency/implementation-manual

[9] 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.

[10] The comprehensive Fiscal Year (FY) 2020-2021 evaluation approach is described in the evaluation plan the team developed in January 2020: https://www.bpa.gov/-/media/Aep/energy-efficiency/evaluation-projects-studies/BPA-2020-21-impact-evaluation-plan.pdf

[11] The FY2020-21 plan also included strategic energy management (SEM) evaluation and assessment of BPA Qualified Commercial HVAC measure savings.

[12] The update to the evaluation plan was completed in December 2022:https://www.bpa.gov/-/media/Aep/energy-efficiency/evaluation-projects-studies/2023-2024-bpa-ee-evaluation-strategy-presentation.pdf

2.3 EVALUATION OBJECTIVES

This document presents the second set of results from the evaluation plan (an impact evaluation of BPA's custom portfolio for industrial customers of Option 2 utilities) conducted by the Evergreen team. The team completed the first set of results for BPA's custom portfolio for industrial customers of Option 1 utilities in July of 2022.[13] The evaluation objectives for this study were to:

1. **Estimate first-year kilowatt-hour (kWh) savings** for the custom industrial portfolio for Option 2 utilities, 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 *Measurement and Verification Protocol Selection Guide* and applying the protocol called *Engineering Calculations with Verification* (ECwV).

[13] See: https://www.bpa.gov/-/media/Aep/energy-efficiency/evaluation-projects-studies/2020-2021-custom-industrial-impact-evaluation-for-option1-utilities-final-report.pdf

3 METHODOLOGY

This section summarizes the methods used to conduct this evaluation and is organized by the following sections: Sample Design, Data Collection, Custom Measure Analysis, and Study and Domain Analysis. Appendix D provides additional detail on the study methods.

3.1 SAMPLE DESIGN

Table 2 shows the number of measures and savings associated with Option 2 utilities' custom industrial measures completed during the study period, by end use and strata. It also includes the study sample allocation of 22 measures (at 19 sites), which includes 19 certainty measures and a stratified random sample of 3 additional measures.

Table 2: Option 2 custom industrial sample design

End Use	Strata*	Reported Savings (kWh)		Number of Measures	Sample Size
		Average	Total	Count	Count
Compressed Air	0	4,861	4,861	1	0
	1	54,597	218,389	4	2
	Certainty	999,802	3,999,206	4	4
HVAC	0	13,087	13,087	1	0
	Certainty	214,728	644,184	3	3
Motors/Drives	0	3,807	3,807	1	0
	1	61,590	123,180	2	1
	Certainty	152,301	304,601	2	2
Process Loads	Certainty	562,517	2,812,584	5	5
Refrigeration	0	18,423	18,423	1	0
	Certainty	107,277	429,107	4	4
Water Heating	Certainty	326,635	326,635	1	1
	Total	**307,465**	**8,898,065**	**29**	**22**

* Stratum *0* denotes the excluded measures (based on very small savings). The *certainty* measures represent a significant portion of total reported energy savings within the end use and are necessary for the evaluation and therefore are not subject to random selection. Process Loads are all assigned to the certainty stratum due to the high variability and small number of measures above and below the size threshold of 200,000 kWh.

** Five refrigeration measures from one utility were removed in October 2021 after being identified as commercial sites that had been mistakenly assigned industrial measure codes. These were replaced with a supplemental sample comprised of measures from the utility between 2019 Q4 and 2020 Q1. The supplemental sample frame added three refrigeration and two compressed air measures.

3.2 DATA COLLECTION

Our general approach to evaluation data collection was to fully leverage the data collected by BPA, project engineers, and the utility program staff throughout the process of developing each project and to only collect additional data from end users if needed to achieve reliable estimates of savings for the sampled measures. We collected the necessary data using a combination of the following approaches (each of which is described in more detail in Appendix D):

- File review
- Telephone/email discussion with project engineers
- Telephone/email discussion with end users
- Site visits
- Affected system trend metering
- Supplemental weather data gathering
- Cost effectiveness parameter data collection

3.3 CUSTOM MEASURE ANALYSIS

We estimated savings for sampled custom measures using the following steps (each of which is described in more detail in Appendix D):

1. Select reliable evaluation model
2. Assess determinant reliability
3. Collect supplemental data
4. Run evaluation model
5. Estimate savings using engineering calculations with verification (ECwV)

3.4 STUDY AND DOMAIN ANALYSIS

Once data collection and analysis were completed for the sample, we compiled a workbook containing all individual site level findings about key drivers for deviations between evaluated savings and original savings estimates. The site-level results were used to estimate the electric savings and cost-effectiveness for the domain and individual end uses within the domain using a ratio analysis. For custom measures with ECwV protocol savings estimates, we also determined when the ECwV protocol provides a reliable estimate of savings. The details of this approach are presented in Appendix D.

4 FINDINGS

This section presents impact evaluation results for BPA's custom portfolio for industrial customers of Option 2 utilities. These findings address the second domain of a rolling evaluation plan with additional domains that will address the entire custom measure and C/I/Ag lighting portfolios (which include all nonresidential lighting and custom measures). The first domain addressed BPA's custom portfolio for industrial customers of Option 1 utilities and was completed in July.

The section is organized as follows:

- Overall results for the second FY2020-2021 impact evaluation domain (Custom Industrial for Option 2 utility customers)
- Project measure level results
- Key drivers of savings
- Lifetime savings
- ECwV savings adjustments
- Cost-effectiveness

Appendix B provides site-specific savings estimation details.

4.1 OVERALL DOMAIN 2 RESULTS

This subsection provides the overall results for this impact evaluation of custom measures installed by Option 2 utility industrial customers with completed reporting between Q2 2020 and Q1 2021.[14]

The overall results showed evaluated savings estimates for Option 2 custom industrial measures as 102 percent of the savings that BPA reported. Evaluators observed that utilities generally followed BPA *Implementation Manual* and M&V protocols correctly and that the difference in realization rates did not result from deviations in procedure.

FIRST-YEAR SAVINGS

Evaluated savings were lower than reported savings for refrigeration and water heating, while the estimated savings for process loads and HVAC were higher than reported. Compressed air and motors/drives were close to the same as reported savings. The difference in evaluated and reported savings in refrigeration and process loads were the most impactful, as shown in Figure 1 and Table 3 (both of which report on the evaluation results for the entire domain).

[14] With a small supplemental sample of measures going back to 2019 Q4. See sampling methodology for more details.

Figure 1: Reported first-year savings by end use compared to evaluated savings by end use

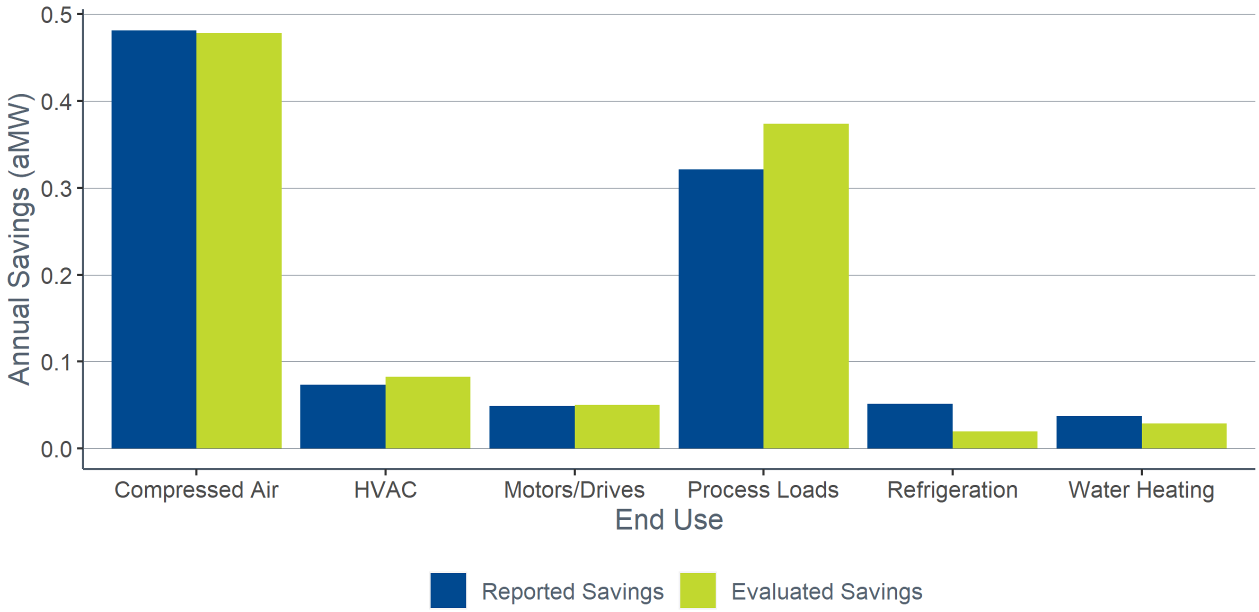


Table 3: Evaluated first-year savings by end use

End Use	Realization Rate	Sampling Relative Precision (90% two-tailed)	Reported Savings (kWh)	Evaluation Savings	
				kWh	Percent of Domain
Compressed Air	99%	5%	4,217,596	4,190,554	46%
HVAC	112%	0%	644,183	723,972	8%
Motors/Drives	103%	33%	427,781	439,983	5%
Process Loads	116%	0%	2,812,583	3,275,749	36%
Refrigeration	38%	11%	447,530	170,813	2%
Water Heating	77%	0%	326,635	251,755	3%
Total	**102%**	**2%**	**8,876,308**	**9,052,828**	**100%**

The realization rate estimated by the prior evaluation of this domain was 107 percent.[15] Note that the evaluation methodology and program cycle differed over the two evaluation periods.[16]

The actual sampling relative precision came in at 2 percent for a 90 percent two-tailed confidence interval. This precision is slightly better than predicted during the sample design phase (10 percent precision). The sample design was based on an expectation of slightly higher variability than observed in the last evaluation of this domain, but the actual variability came in substantially lower.

COVID IMPACTS ON EVALUATED SAVINGS

To evaluate COVID-19 impacts on evaluated savings, the evaluators collected information on each sampled site regarding any adjustments to facility operation (e.g., changes in production schedules). Evaluators calculated a second set of savings estimates that are intended to represent what would have happened had COVID-19 not occurred. Those results were very similar to the evaluated savings with no adjustments for COVID-19, therefore, the evaluation results reported in this report's main body are the evaluated savings with no COVID-19 adjustments. See Appendix C for the detailed comparison.

COMPARISON TO DOMAIN 1 OVERALL RESULTS

The domain 2 realization rate (102%) was substantially higher than the realization rate for domain 1 (85%), as shown in Table 4.[17] The Option 1 study had more projects with overestimates of savings than domain 2 for this particular time period. Additionally, the Option 1 study included a finding that reduced savings for projects that claimed evaporative cooling as part of their savings (a difference in baseline assumptions).

---

[15] See results for the Option 2 Non-Lighting Industrial domain from the prior evaluation completed in 2017: https://www.bpa.gov/-/media/Aep/energy-efficiency/evaluation-projects-studies/impact-evaluation-site-specific-portfolio-final-report.pdf.

[16] In the prior study, Option 2 utilities had the opportunity to fund additional sample points to create their own customized evaluations. These sample points enhanced the precision of the BPA study. The utilities were given the opportunity to review new models after BPA review. In this study, Option 2 utilities did not fund oversamples and BPA only reviewed new models internally.

[17] See results for the Option 1 Custom Industrial impact evaluation: https://www.bpa.gov/-/media/Aep/energy-efficiency/evaluation-projects-studies/2020-2021-custom-industrial-impact-evaluation-for-option1-utilities-final-report.pdf.

Table 4: Comparison of Domains 1 and 2 Realization Rates

Domain	Realization Rate	Sampling Relative Precision (90% two-tailed)	Reported Savings (kWh)	Evaluation Savings (kWh)
Option 1 Custom Industrial	85%	7%	38,944,452	33,230,310
Option 2 Custom Industrial	**102%**	**2%**	**8,876,308**	**9,052,828**

4.2 MEASURE LEVEL RESULTS

Results at the project measure level based on the evaluation sample of 22 projects were highly variable, with realization rates ranging from -0.2 to 1.8, as shown in Figure 2, with each project measure in the sample represented by a single point. The points are grouped by end use and size stratum (different colors for each measure category, a circle for certainty sites and a square for other size stratum), and the size of the shapes corresponding to the size of evaluated savings for each stratum. For water heating (in yellow) and refrigeration measures (in teal blue), there were more low realization rate project measures than high realization rate project measures, leading to low overall realization rates at the end use level. For all other end uses, realization rates were relatively variable, with some projects well above 1.0 and others well below 1.0, leading to realization rates close to 1.0.

Figure 2: Project measure-level realization rates

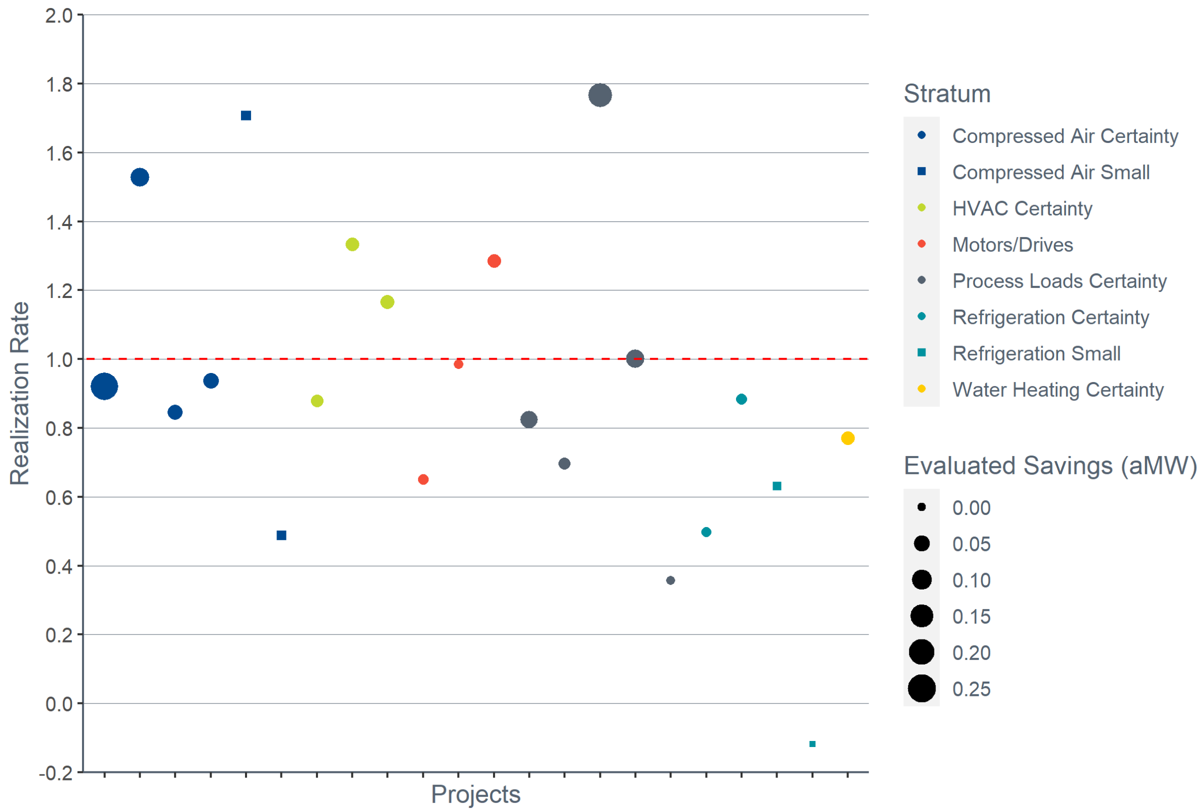
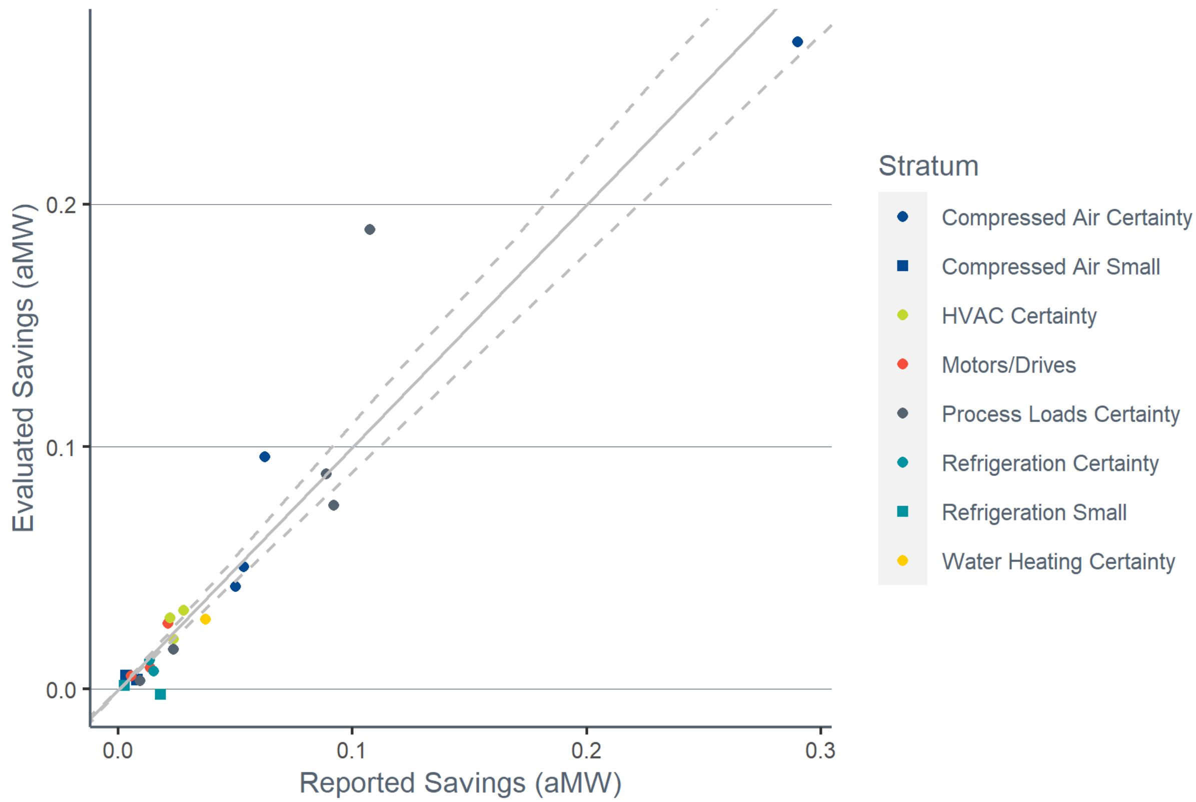


Figure 3 shows evaluated savings results by project measure for the evaluation sample, expressed in average megawatts (aMW). Points lying above the gray diagonal line represent end uses with evaluated savings higher than reported savings, while those lying below the gray diagonal line represent end uses with evaluated savings lower than reported savings. The dashed lines indicate +/- ten percent of reported savings.

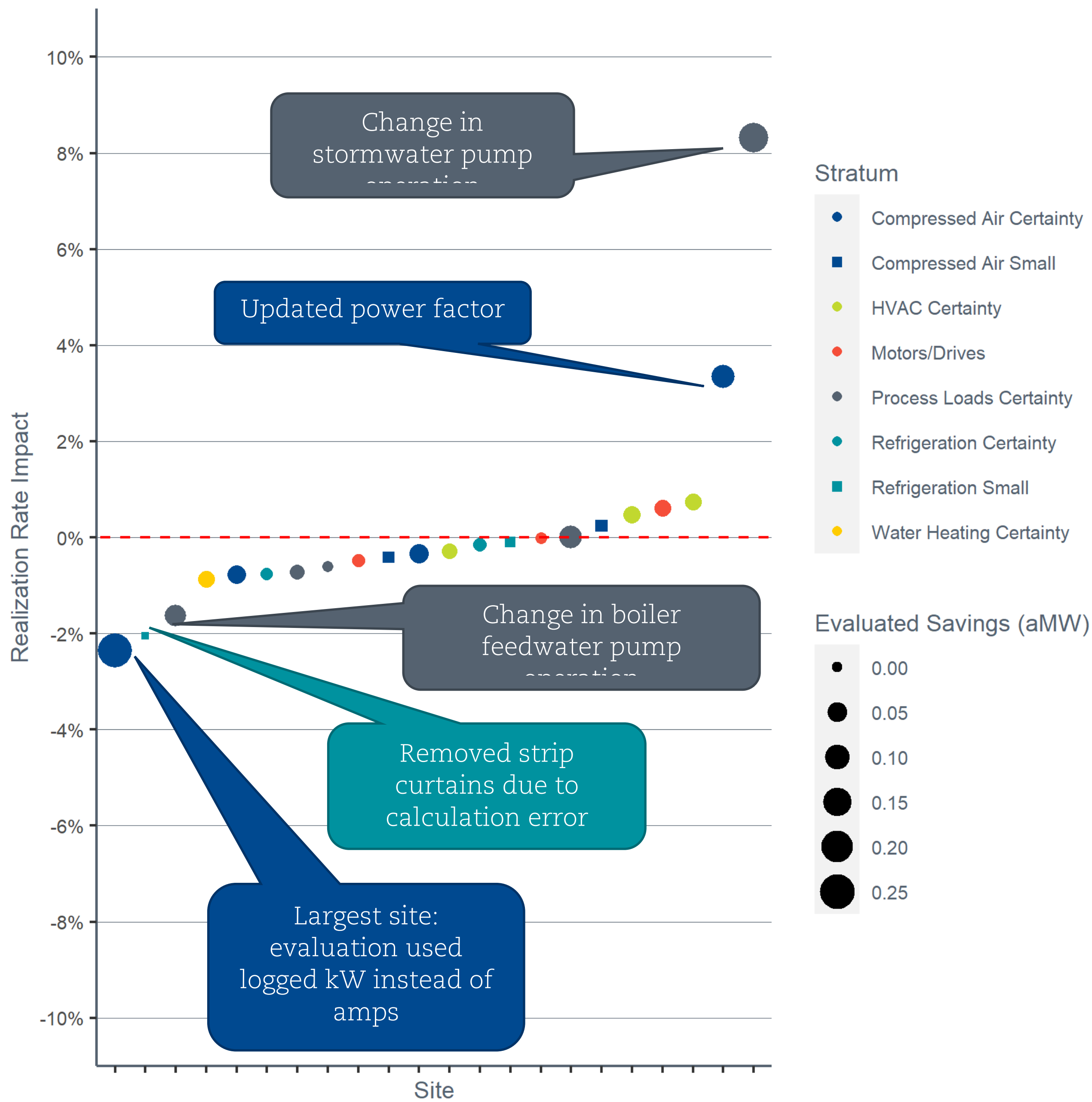
Figure 3: Evaluated savings versus reported savings by project measure



4.3 KEY DRIVERS OF SAVINGS DIFFERENCES

Figure 4 shows the impact of each reviewed project measure in the evaluation sample on the overall domain realization rate. Project measures below the red dashed line are driving the domain realization rate below 1.0, while project measures above the line are driving the domain realization rate above 1.0. Most projects had little or no influence on the overall realization rate either because their realization rates were near 1.0 or their small size made them noninfluential on the total. While most projects had realization rates near or below 1.0, a small number of large projects had realization rates well above 1.0, increasing the overall realization rate above 1.0. The callout boxes within the figure summarize the reasons for some of the most influential projects on the deviation in the realization rate from 1.0, which are discussed in more detail after the figure.

Figure 4: Deviations in domain-level realization rates



The most influential project measures **negatively** impacting the domain-level realization rate had the following issues:

1. **Evaluation used logged kW instead of amps.** This is the largest project in the sample, so even a minor difference leading to the 92% realization rate will have an oversized impact on this chart. The evaluators chose to use manufacturer's reported power for one piece of equipment instead of power calculated from reported amps because it was consistent with the method used for all other equipment in the implementer's methodology. True power could not be safely verified due to the equipment being

medium voltage. This result on its own appears to be a unique occurrence with oversized impact and does not merit systematic correction.

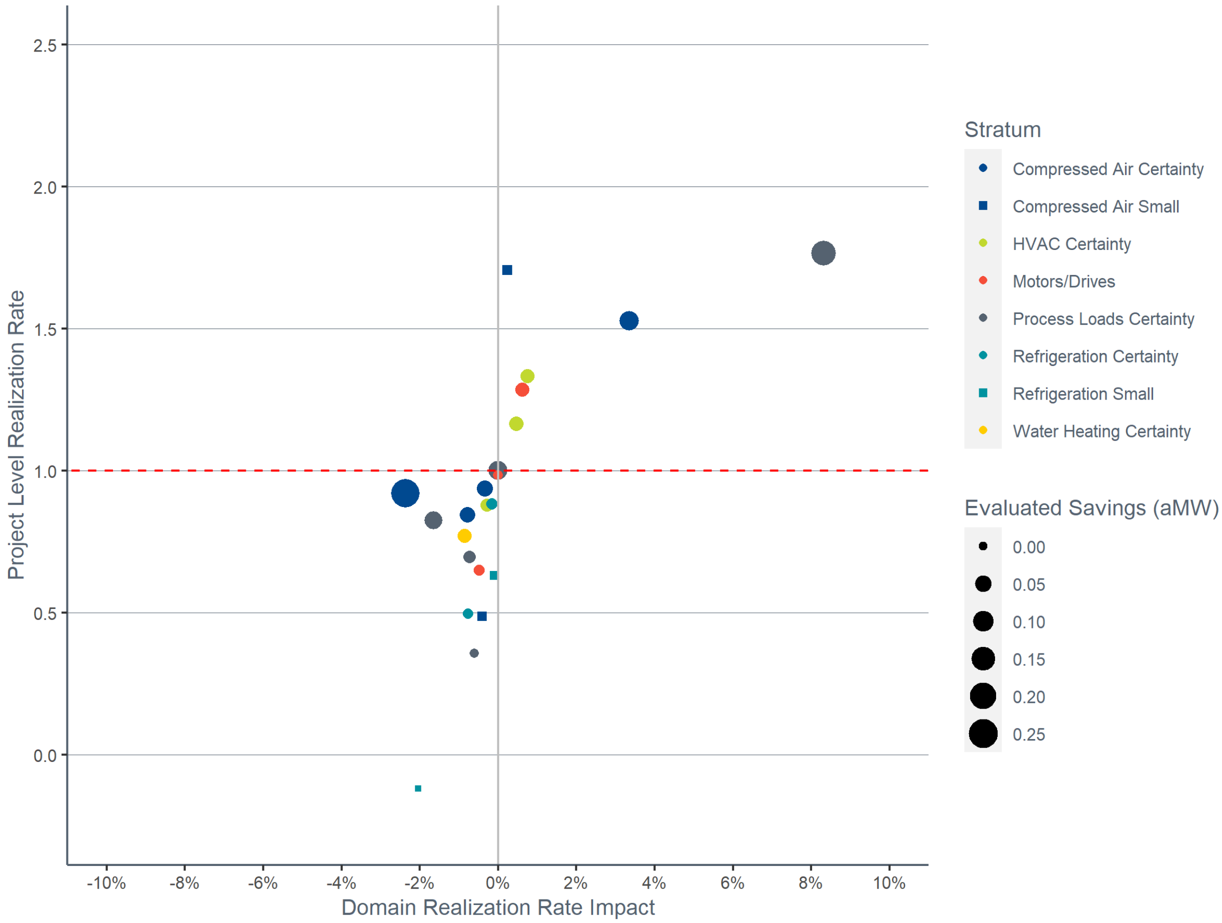
2. **Removal of strip curtains, due to the use of non-standard assumptions for estimating impacts of strip curtains.** One small refrigeration project removed strip curtains as part of an automated door closer project. The typical regional assumptions for strip curtains value their savings higher than the calculator used for this project. This caused a significant increase in energy consumption, negating all other savings associated with the project.
3. **Differences in observed operating conditions**. Some projects had different observed conditions for the first year of operation than what was documented in the original M&V period.
    a. The evaporator fan duty during implementation was overestimated for a very large refrigeration project.
    b. An M&V production value was used for boiler feedwater that was substantially lower than was found from two years of post-implementation data for a large process project. Since data varied, two years of post implementation data were used to create a "typical first year of operation."

The most influential project measures **positively** impacting the domain-level realization rate had the following issues:

1. **Use of a fixed power factor.** A fixed power factor of 0.9 was used for one project instead of calculating as a percentage of full load amps, and CAGI data was assumed instead of actual operation.
2. **Use of a lower deemed savings value.** A deemed savings load profile was used for one project that was lower than collected use data.
3. **Differences in observed operating conditions**. Observed conditions were used for some projects that were different for the first year of operation than what was documented in the original M&V period.
    a. An M&V value was used for influent MGD flow for a large process project that was substantially higher than was found from two years of post-implementation data.
    b. Substantially higher rates of production were observed for two sites.

The project measure impact map in Figure 5 shows the relationship between realization rate, size of project, and resulting overall impact on the domain realization rate based on the evaluation sample results. The project impact map combines all the information presented in the previous series of exhibits. The x-axis shows the impact on the overall realization rate, while the y-axis shows the project measure realization rate. Project measures in the lower left quadrant are driving the domain realization rate below 1.0, while project measures in the upper right quadrant are driving the domain realization rate above 1.0. Larger dots represent larger project measures, which generally increase the total impact on realized energy savings.

Figure 5: Project measure impact map



4.4 LIFETIME SAVINGS

Table 5 on the next page shows the estimated evaluated lifetime savings for the sampled sites. The evaluated lifetime savings estimates are slightly higher than the reported lifetime savings.

Table 5: Lifetime savings

End Use	Evaluated Lifetime Savings (kWh)	Reported Lifetime Savings (kWh)	Lifetime Savings Realization Rate
Compressed Air	40,716,516	40,979,254	99%
HVAC	8,149,282	7,216,925	113%
Motors/Drives	3,634,367	3,533,570	103%
Process Loads	32,757,494	28,125,830	116%
Refrigeration	2,320,483	5,637,405	41%
Water Heating	3,776,328	4,899,525	77%
Total	**91,354,469**	**90,392,510**	**101%**

Some minor adjustments to lifetime savings were made by changing the TAPs and their corresponding lifetimes for evaluated measures including changing: [18]

- HVAC Control Improvements (variable frequency drive [VFD]) (5-year lifetime) to Motors/Drives Installation on Fan System (10 years);
- HVAC Control Improvements (VFD) (5-year lifetime) to Motors/Drives Control Improvements (VFD) (10 years); and
- Interactive Refrigeration System Improvements (10-year lifetime) to Packaged Refrigeration System Improvements (15 years).

However, the net impact of lifetime adjustments was very small compared to the adjustments in first-year savings. The resulting lifetime savings realization rate of 101 percent is slightly lower than the first-year savings realization rate (102%).

4.5 ECWV SAVINGS ESTIMATION EFFICACY

The evaluation looked at two issues with *Engineering Calculations with Verification* (ECwV) savings estimation methods:

1. Can ECwV provide an unbiased estimate of energy savings?
2. Did the Evaluator's implementation of BPA's ECwV M&V protocol provide an unbiased estimate of energy savings?

For the first question, the evaluation team estimated savings using an ECwV methodology and a regular high rigor evaluation method for all projects in the sample. The results show that ECwV approaches as applied by the evaluation team provided unbiased results for smaller projects while underestimating savings for larger projects, as shown in the stratum level results in Figure 6 below.

[18] TAP stands for Technology/Activity/Practice. The BPA reporting system uses a standardized taxonomy (TAP) for classifying measures. For most projects, BPA or utility staff assigns one of 86 TAP descriptions to each physical measure or change implemented as part of a project. In cases where the tracked TAP was found to be improperly assigned, evaluators corrected the TAP for purposes of estimating lifetime.

Figure 6: Evaluated savings with and without ECwV

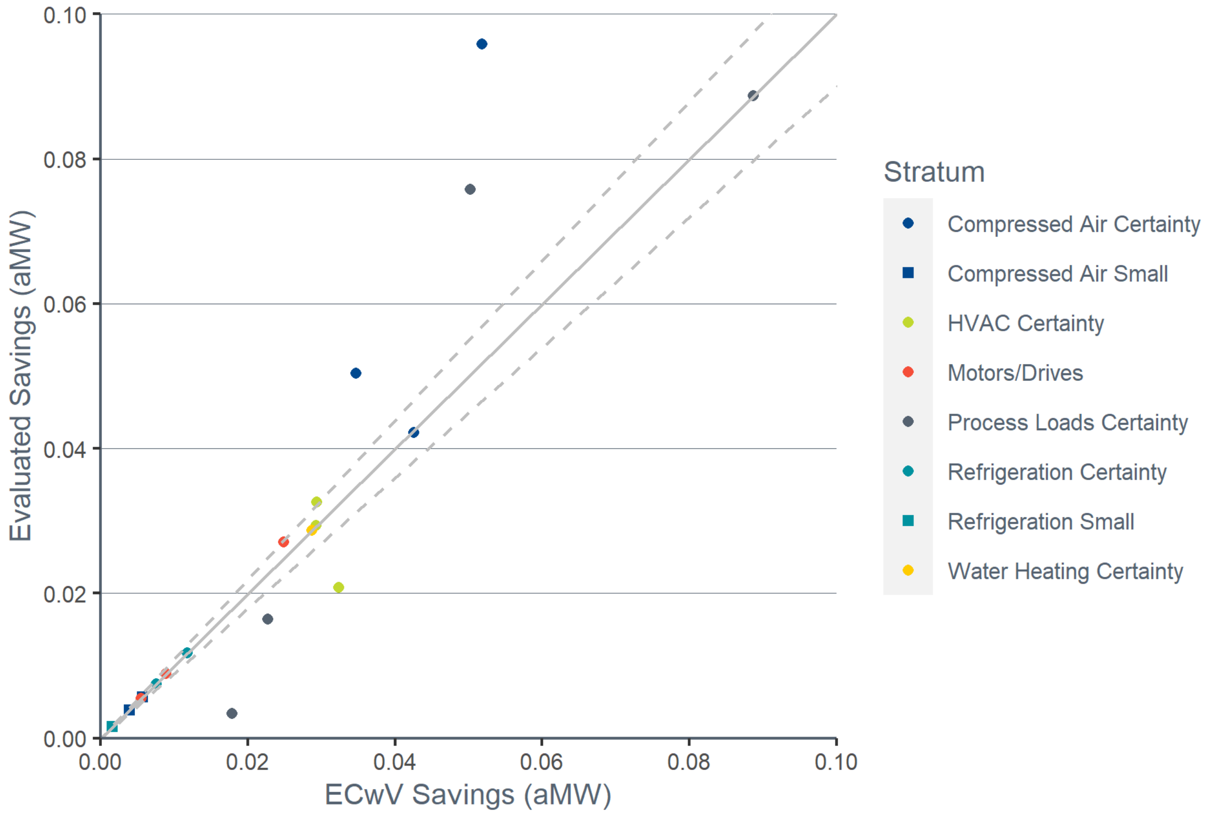


Table 6 compares evaluated savings with ECwV methodology applied to evaluated savings without at the stratum level. The total savings of all projects in each stratum are shown in the table; average savings per project are lower. If the ratio of the two is near 100 percent, it means that ECwV provided an unbiased estimate of the true savings. The small strata all have ratios near 100 percent, while the medium and large strata are more variable, with many ratios below 100 percent. This means that ECwV was relatively accurate for small projects but substantially underestimated savings for many large projects.

Table 6: Evaluated savings with and without ECwV by end use and size

End Use	Size Range	Evaluated ECwV Stratum Savings	Evaluated Stratum Savings	Ratio	Sample Size	Population
Compressed Air	Small (<200,000 reported kWh / project)	83,669	83,669	100%	2	4
Compressed Air	Large	4,455,548	3,987,983	112%	4	4
HVAC	Large	797,740	723,972	110%	3	3
Motors/ Drives	All	344,337	363,437	95%	3	4
Process Loads	Large	3,111,247	3,275,749	95%	5	5
Refrigeration	Small (<200,000 reported kWh / project)	-5,155	-5,155	100%	2	2
Refrigeration	Large	168,936	168,936	100%	2	2
Water Heating	Large	251,755	251,755	100%	1	1
Total		**9,208,078**	**8,850,346**	**104%**	**22**	**26**

4.6 COST-EFFECTIVENESS

The domain is strongly cost-effective overall, based on the evaluation results, producing $4.89 in benefits for every dollar spent. The evaluated benefit-cost ratio (4.89) came in slightly higher than BPA's reported benefit-cost ratio (4.78). All end uses have both reported and evaluated benefits greater than costs, as shown in Table 7.

Table 7: Benefit-cost results

End Use	Reported			Evaluated		
	Benefits ($)	Costs ($)	Benefit-Cost Ratio	Benefits ($)	Costs($)	Benefit-Cost Ratio
Compressed Air	$3,010,867	$622,875	4.83	$2,991,563	$622,875	4.80
HVAC	$473,301	$112,582	4.20	$531,924	$112,582	4.72
Motors/Drives	$259,622	$54,273	4.78	$267,028	$54,273	4.92
Process Loads	$2,066,488	$426,055	4.85	$2,406,790	$426,055	5.65
Refrigeration	$315,294	$110,646	2.85	$120,335	$110,646	1.09
Water Heating	$239,988	$4,359	55.05	$184,972	$4,359	42.43
Total	**$6,365,560**	**$1,330,789**	**4.78**	**$6,502,611**	**$1,330,789**	**4.89**

5 KEY FINDINGS AND RECOMMENDATIONS

The evaluation team developed recommendations and considerations resulting from the key findings of this evaluation.

Key Finding: The overall realization rate for the custom portfolio for industrial customers of Option 2 utilities was 102%, yet the variability of realization rates was relatively high. The overall realization rate of 102% indicates that on average, claimed savings are equal to realized savings. Yet, more than half of projects had realization rates either below 80% or above 120%. On an individual project basis, reliability is a factor for utilities and end-use customers to ensure they are achieving savings and paying incentives as intended. Therefore, there is room to improve the reliability of individual project savings. Recommendations 1 through 4 offer specific suggestions that, if taken, should lead to closer alignment of reported savings with evaluated savings for individual projects.

Key Finding: Evaluators observed that BPA *Implementation Manual* and M&V protocols were generally followed correctly, except that savings were not always estimated for the first year post implementation. For some evaluated projects, the difference in reported savings and evaluated savings is due to savings not being estimated for the first year of operation after the measure is installed.

> **Recommendation 1: BPA should clarify requirements for the basis year of savings in the BPA *Implementation Manual*.** The implementation manual should clearly state what first year savings means and what the basis year is for estimating savings. E.g., "*The first year of savings should be calculated assuming that the post implementation period represents the new norm of operation. If there have been changes in standard operating procedures, then the baseline model should be adjusted to match conditions for the first year of operations.*"

Key Finding: The evaluators observed some energy models that were not consistent with regional Custom Project practices. This included:

- A refrigeration model using strip curtain insulation values different from ASHRAE, RTF deemed, and BPA UES measures of a similar kind.
- A compressed air model with deemed savings values based on size.

> **Recommendation 2: BPA should offer training and access for Option 2 utilities to BPA's solutions for common measures.** Access to BPA's library of approved solutions may reduce engineering overhead for utilities and help to develop more consistent regional practices. This can include solutions and models that are fully IPMVP compliant or simplified approaches for common measures (NWRCAT, VFD calculators, etc.).

Key Finding: Small and medium-sized projects showed similar results in evaluated savings using the BPA *Engineering Calculations with Verification* (ECwV) protocol or high-rigor M&V methods. While there was variability from site to site and the ECwV results are less precise, ECwV results were similar to evaluated results overall. This finding is consistent with the results from a similar analysis conducted for this program for Option 1 utilities.

> **Recommendation 3: BPA should consider applying ECwV to a wider size range of projects.** While the evaluation sample size is not large enough to provide a suggested size level, there is some evidence that the savings threshold for ECwV could be increased from what the ECwV guide recommends (200,000 kWh). However, at least anecdotally, projects with more interactive effects between different pieces of equipment tend to have lower

accuracy ECwV results, so it makes sense to add to the M&V protocol selection guide a brief assessment of the level of risk in savings estimation error due to the simpler protocol with fewer verification requirements. More ECwV projects may reduce overall engineering and administrative load for energy efficiency programs and liberate resources for more energy efficiency projects.

Recommendation 4: BPA should conduct outreach to Option 2 utilities to promote the use of ECwV as a vetted M&V approach for smaller projects that still fully complies with BPA M&V guidelines.

Key Finding: **COVID-19 did not have a substantive impact on evaluated savings** for either domains 1 or 2 of the custom industrial program, based on evaluation data collection conducted over the last two years.

Consideration: Unless the response to the COVID pandemic changes substantially, it is not necessary to spend evaluation resources collecting a second set of results that attempt to factor in the impacts of COVID-19 on realized savings.

Key Finding: For the most recent Option 1 and Option 2 custom industrial evaluations, the evaluators stratified each study sample by end use and size of project. The variation in realization rates did not differ based on end use, and the realized precision estimates for both studies exceeded the targeted 90/10 confidence level.

Consideration: For future custom industrial evaluations, BPA should consider stratifying the sample by Option 1 versus Option 2 utility and size of project, and dropping the end use stratification. Based on the sample precision realized for both studies, the sample size could also be reduced going forward.

APPENDICES

APPENDIX A: DEFINITION OF KEY TERMS

We rely on the following definitions of key terms throughout this report.

Reporting System

BPA uses its reporting system to track projects completed by public power utilities under various programs and initiatives. For Option 1 utilities, BPA has detailed custom project proposals and completion reports in its system (Option 1 Custom Project Calculator). Option 2 utilities report high-level project information through a custom project completion report into the BPA system periodically (Option 2 Custom Project Calculator).

Domain

Domains are components of the portfolio. They are defined by Option (utilities are Option 1 or 2 for measurement and verification [M&V] purposes), Measure Type (e.g., Lighting, Non-Lighting), and Sector (Industrial and Commercial or the combination of commercial and agricultural for Option 1 utilities).

Option 1

For Option 1 utilities, BPA manages the bundle of energy savings from custom projects. This requires that BPA manage the portfolio risk for both project performance and cost-effectiveness. Often, BPA is involved throughout the project lifecycle by providing technical support, M&V implementation, approval of projects and oversight/evaluation.

Option 2

For Option 2 utilities, the customers manage the bundle of savings from their custom projects. This entails the customers managing the risk of project performance and cost-effectiveness by conducting all aspects of M&V and custom project quality control (e.g., project proposal and project completion documentation review) internally.

Project

A project is a phase of work at an end user location that improves energy efficiency. An end user is the customer of a BPA utility. The project tracking data record a date when the project is complete. The data also contain information such as the name of the end user, the location where the work was carried out, and other data critical to this evaluation. End users may authorize the completion of many phases of work, each of which is tracked as a separate project in the BPA reporting system.

Measure

A measure is a distinct Technology/Activity/Practice (TAP) within a project. The BPA reporting system uses a standardized taxonomy (Technology/Activity/Practice) for classifying measures. For most projects, BPA or utility staff assign one of 86 TAP descriptions to each physical measure or change implemented as part of a project.

Project Engineers

Project engineers assist in the identification, development, savings estimation, cost-effectiveness analysis, M&V, and quality control review of projects. Project engineers may be BPA staff, utility staff, or staff of BPA or utility project implementation contractors. For the purposes of this evaluation, project engineers are not staff or contractors employed by the end users, even though the end user workforce may have played an important role in the development of a project.

M&V Model

This M&V model (an algorithm or calculation procedure) is the model used by project engineers to estimate savings for the measures that comprise a project. The NWRCAT compressed air calculator is an example of such a model. Models for other measures might be building simulation models such as eQUEST, custom-engineered or standardized spreadsheet calculators, and custom regression models (such as those developed using Energy Performance and Carbon Emissions Assessment and Monitoring [ECAM]).

Reported Savings

Reported savings are the savings estimated by the project engineers and entered in the BPA reporting system. These savings are based on the M&V model. Please note that the BPA system uses the term “estimated savings” for the savings estimated at the proposal stage and “actual savings” for the savings at the completion report stage. The BPA *Implementation Manual* does not require all projects to submit a formal proposal. Reported savings are based on the “actual savings” field in the reporting system. “Actual savings” is busbar savings, equal to 1.09056 times site savings.

Evaluation Model

This is the model selected by our evaluation team to re-estimate savings for sampled measures. The same types of models as listed above for the M&V models were also options under consideration. Please note that although the evaluation model may differ from the M&V model, this does not necessarily mean that the M&V model was inappropriate for the project. Therefore, there may be cases where a more reliable model is used in evaluation of a sampled measure, even though that model would not be cost-effective for M&V on all measures.

Evaluation Savings (and Evaluated as-observed versus Evaluated without COVID-19-Specific Terminology)

Evaluation savings are the savings estimated by the evaluation team. These savings are based on the evaluation model and rely on best practical data collection and savings estimation practices, as laid out in the RTF guidelines and informed by evaluator experience. The evaluation estimated the savings achieved during the first year of measure operation. If any of the evaluation data collection occurs more than one year after the measure was complete, it may indicate failures in the measure performance that are relevant to measure lifetime and not to the first-year savings. Evaluation savings estimates reflect the conditions of the measure during the first year of operation.

For this evaluation where the project measure operations under evaluation occurred during the COVID-19 pandemic, the evaluators estimated two separate sets of results:

1. Evaluated savings without Covid: the definition provided above for evaluated savings, with any changes to site operations resulting from COVID-19 reported by the site respondents removed. Site-specific findings may also be referred to as as-expected savings.
2. Evaluated savings: the traditional definition of evaluated savings based on actual observations, with no adjustments for COVID-19. Site-specific findings may also be referred to as as-operated, or as-found, savings.

Current Practice Baseline

BPA and the Regional Technical Forum (RTF) have different definitions of current practice baseline for custom measures, based on BPA's M&V selection guide[19] and RTF guidelines,[20] respectively.

BPA: "*When the practitioner uses a current practice baseline, the efficiency level of the baseline equipment must be consistent with any state or local mandates for new equipment, which may vary from city to city and state to state.*"

RTF: "*The practitioner needs to identify what would normally be done, based on prior experience with similar projects. The practitioner should start by using applicable codes and standards, or one of the following if they constitute a more energy efficient baseline for the measure and the information is practical to obtain and applicable to the delivered measure's location OR there is no applicable code or standard for the measure implemented.*

- *Recent similar purchases by the end user*
- *Documented end user plans or specifications*
- *End user or vendor developed alternative designs, considered as part of the measure selection process*
- *End user description of what was done in similar circumstances elsewhere in the facility or in another facility they operate*
- *Equipment vendor's description of what they would normally do for this end user*"

Realization Rate

Realization rate is the ratio of evaluation savings to reported savings. Realization rates greater than 1.0 mean that we found more savings than were reported.

[19] BPA. 2018. *Measurement & Verification (M&V) Protocol Selection Guide and Example M&V Plan*. Page 7. https://www.bpa.gov/-/media/Aep/energy-efficiency/measurement-verification/1-bpa-mv-selection-guide.pdf (file will download automatically when clicking)
[20] Regional Technical Forum. 2020. *Guidelines for the Assessment of Energy Efficiency Measures*, Section 4.3.3, page 20. https://nwcouncil.app.box.com/v/2020RTFGuidelines

Key Determinants

Key determinants influence the savings from a measure. The evaluation considered the following key determinants:

- **Connected load.** Baseline or efficient-case rated kW demand and/or the quantity of the equipment.
- **Efficiency profile.** Part-load impacts on demand profile, including variable frequency drives (VFDs) and HVAC interaction factors.
- **Hours of operation.** Baseline or efficient-case schedule of operation for a measure.
- **Load profile.** Facility occupancy rates and changes not captured by other categories of key determinants.
- **Production.** Number of production units per unit of time.
- **Weather.** Weather-based data used for weather-sensitive measures, such as dry and wet-bulb temperatures, or heating and cooling degree-days.

Reasons for Difference

The reasons for difference are what we changed that caused a modification to one or more key determinants and ultimately savings. We ranked the impacts as primary or secondary increase/decrease to give a sense of their scale. We assigned all reasons to one of the following categories:

- **Documentation error.** These include errors in calculations or values entered into models.
- **Other.** Commonly, a change in inputs due to a contradicting finding in the first year. This would indicate that the value for the key determinant in the project documentation was correct, but the value entered in the savings calculation did not match what was in the documentation. It could also indicate that the key determinant in the project documentation did not match what was found during the site visit or in trend data.

Measure Baseline

Measure savings must be determined against clearly defined baseline conditions. The RTF guidelines define two possible baseline conditions that were used in this evaluation:

- **Current practice.** A current practice baseline is used if the measure affects systems, equipment, or practices that are at the end of their useful life. The baseline is defined by the recent typical choices of the end user in purchasing new equipment and services. Current practice baseline is also used for new construction projects where there is no pre-existing systems, equipment or practices.
- **Pre-conditions.** A pre-conditions baseline is used when the measure-affected equipment or practice still has remaining useful life. The baseline is defined by the existing condition at the end user site just prior to the delivery of the measure.

ProCost Model

ProCost is a spreadsheet tool, developed by the Northwest Power and Conservation Council that computes regional measure lifecycle cost-effectiveness. ProCost uses regional economic and power system assumptions that are updated with each Council Power Plan.

Measure Lifetime

Measure lifetime, according to the RTF guideline for lifetime savings, is defined as the median number of years during which at least half the deliveries of a measure are in place and operable, i.e., producing savings. For example, consider the installation of 100 VFDs on pumps. If the VFDs were regularly inspected for many years, it would be possible to determine when each one became inoperable (failed mechanically or electrically or was removed from service). The lifetime for the measure would be the median number of years to measure failure, i.e., no longer producing savings. An estimate of measure lifetime is a required input to ProCost.

Incremental Costs and Benefits

When a measure is delivered, costs are incurred and benefits realized—e.g., the value of electricity savings and other nonelectric benefits such as changes in operations and maintenance expenses. Only incremental costs and benefits are used in estimating life cycle costs and benefits.

A measure's incremental costs and benefits are those incurred in the efficient case delivery, beyond what is required to establish and maintain the baseline condition. For a precondition baseline, the baseline does not involve any change and thus baseline costs and benefits are zero. In this case, incremental costs and benefits are equal to the efficient case costs and benefits. For measures with a current practice baseline, the baseline condition does require a change and therefore has costs and benefits. In this case, the incremental costs are the difference between the efficient case and the baseline case delivery.

NEBs (Nonelectric Benefits)

Nonelectric benefits are defined as any benefit, positive or negative, that is not captured by the value of the electric savings or the measure incremental cost. NEBs include changes caused by the measure in the costs of operation and maintenance or other utilities such as gas, water or wastewater. Further explanation of these benefits can be found in the RTF guidelines (see the guideline for the estimation of incremental measure costs and benefits).

Total Resource Cost (TRC) Test

The TRC is one type of cost-effectiveness testing that includes all incremental cost and lifetime benefits of a measure, regardless of who pays for or receives them. BPA uses the definition of the TRC test consistent with the Northwest Power and Conservation Council.

APPENDIX B: SITE-SPECIFIC SAVING ESTIMATION

This appendix provides the measure level results for the sample.

Detail ID	Stratum	Site Realization Rate	Site Impact on Domain Realization Rate
101	Compressed Air Certainty	92%	-2%
201	Process Loads Certainty	82%	-2%
501	Compressed Air Certainty	94%	0%
701	Motors/Drives	129%	1%
702	HVAC Certainty	117%	0%
901	Refrigeration Small	-12%	-2%
1201	HVAC Certainty	88%	0%
1501	HVAC Certainty	133%	1%
1701	Compressed Air Small	170%	0%
1901	Compressed Air Small	49%	0%
2001	Motors/Drives	65%	0%
2101	Process Loads Certainty	70%	-1%
2301	Process Loads Certainty	177%	8%
2602	Process Loads Certainty	100%	0%
2603	Compressed Air Certainty	152%	3%
2801	Process Loads Certainty	36%	-1%
2901	Water Heating Certainty	77%	-1%
2902	Motors/Drives	99%	0%
3001	Refrigeration Certainty	50%	-1%
3301	Compressed Air Certainty	84%	-1%
3401	Refrigeration Small	63%	0%
3402	Refrigeration Certainty	88%	0%

APPENDIX C: COVID IMPACTS ON EVALUATED SAVINGS

The evaluators collected information on each sampled site regarding any adjustments to facility operation for COVID-19 (e.g., changes in production schedules). Evaluators calculated a second set of savings estimates that are intended to represent what would have happened had COVID-19 not occurred. Those results were very similar to the evaluated savings with no adjustments for COVID-19. The evaluation results reported in this report's main body are the evaluated savings with no COVID-19 adjustments.

Figure 7 below compares the end use level results with and without COVID-19. Evaluated savings values for with and without COVID are equal in all end uses except Motors/Drives. Within Motors/Drives, only one site has different with and without COVID values.

Figure 7: Reported first-year by end use compared to evaluated savings (with and without COVID) by end use

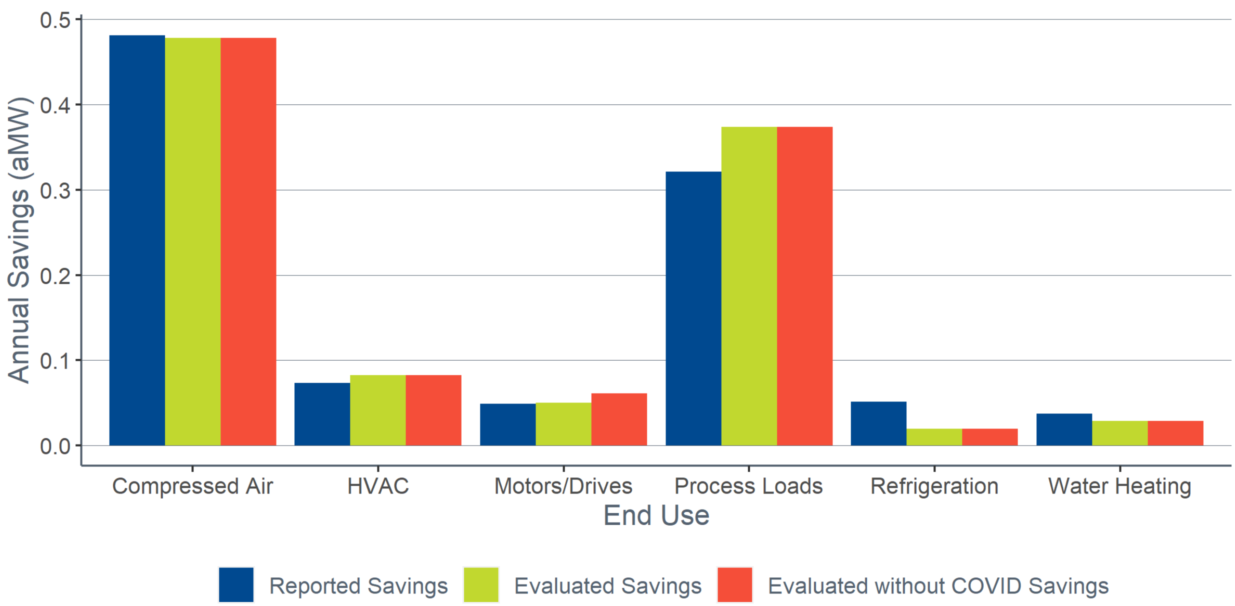
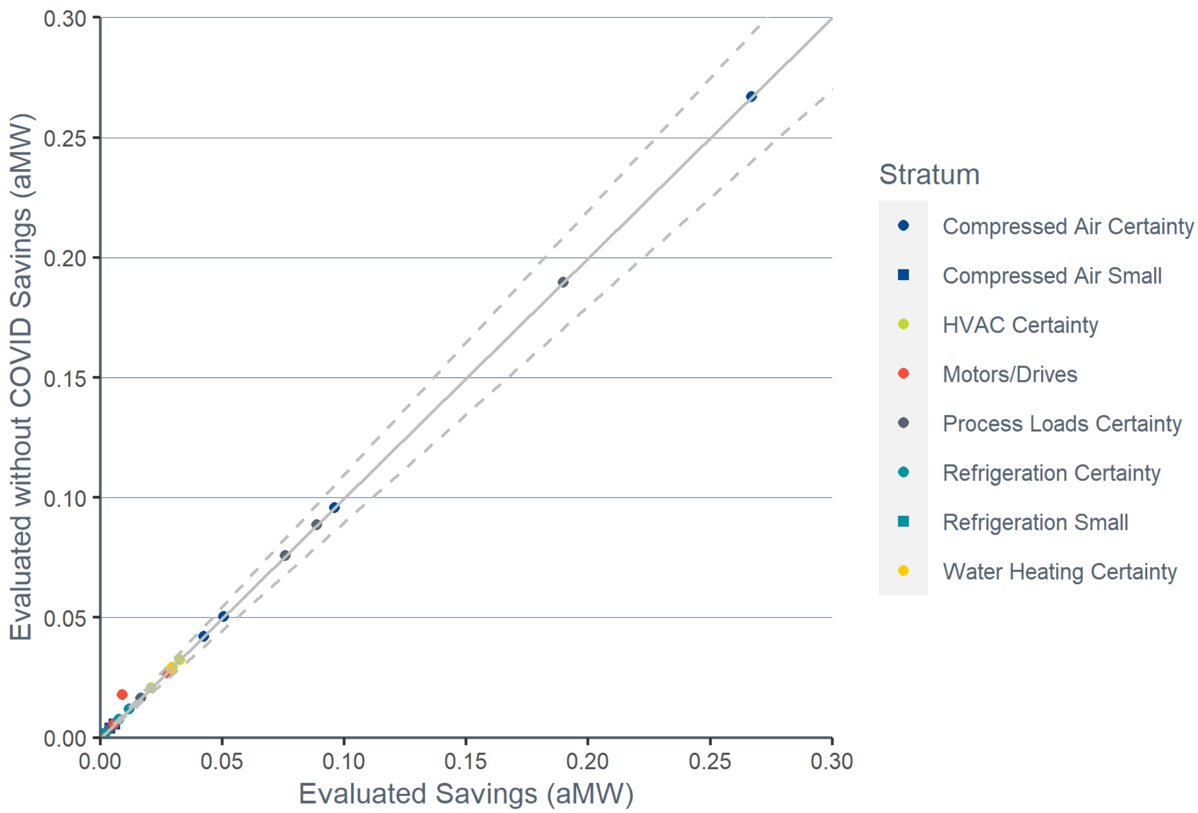


Figure 8 below compares the individual project measure results with and without COVID-19 results included. Points lying above the gray diagonal line had evaluated savings without COVID (i.e., with estimated effects of COVID removed) higher than evaluated (as-observed / with COVID) savings, while those lying below the gray diagonal line had evaluated savings without COVID lower than evaluated savings without COVID. The two dashed gray lines mark +/- 10 percent of as-observed savings. Only one project showed substantial differences. This difference was due to COVID-induced changes in business operations during the measured post-installation period.

Figure 8: Project measure comparison between evaluated results (with and without COVID-19)



APPENDIX D: DETAILED METHODOLOGY

This appendix provides more detail on the study methods (sample design, data collection and analysis).

SAMPLE DESIGN

BPA's evaluation policies have established a target for impact evaluation, striving for evaluations that attain a relative error of 10 percent at the 90 percent confidence level, with a minimum acceptable level of 80/20. The study's sampling strategy targeted a 90/10 confidence level and precision for Domain 2.[21] The sampling unit of this study is a measure, defined as a unique Technology/Activity/Practice (TAP) for a single project at a distinct site (as defined by utility-assigned site ID and facility address).[22]

The Option 2 Domain 2 evaluation was primarily focused on projects with claimed savings that were completed between Q2 2020 and Q1 2021, with a small supplemental sample of measures from one utility going back to 2019 Q4.[23] Our initial sample was based on BPA tracking data (summary BOOM report data), pulled on July 13, 2021; the sample was updated with BPA tracking data (the detailed IS2.0 data) from August 2021, and then revised (IS2.0 data) in November 2021.

The sampling was conducted with a conventional optimum allocation stratified design based on end use category and reported kWh savings for the measure.[24] We defined an excluded stratum (i.e., stratum 0) that contains very small measures; this is the group of measures that collectively account for less than 1 percent of the domain savings. Measures that represent a significant portion (more than 25%) of total reported energy savings for an end use are assigned to a priority "certainty" stratum. We consider these measures necessary for the evaluation; thus, they are not subject to random selection. Moderately sized measures were then allocated to a probabilistic strata, defined by an upper bound of 200,000 kWh savings. Between the probabilistic strata and the certainty strata (above 200,000 kWh), the sample design ensured a mix of measure sizes and ensuring that our sample contains sufficient projects for the ECwV analysis.

DATA COLLECTION

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

---

[21] Actual precision on evaluated savings was 90/2.
[22] For uniformity of evaluation approach, evaluation and project resource management, and cost control, sampling is based on measure.
[23] We started with a 12-month period, from 2020 Q2 to 2021 Q1 and invoice approval dates between June 3, 2020 and June 2, 2021. After the utilities were notified of the customers in the sample, the team identified five refrigeration measures from one utility had been installed at commercial facilities and then been mistakenly assigned industrial measure codes. The sites were removed from the sample frame and replaced with measures between 2019 Q4 and 2020 Q1. For the end use strata (and three of the four utilities), the evaluation cycle still spans 12 months. However, due to this issue, the timeline for the supplemental sample differs from the rest.
[24] In BPA taxonomy, TAPs roll up into end-use groups. Therefore, where feasible, the evaluation will attempt to roll up results into end uses for additional insight to BPA. There is insufficient sample to achieve 90/10 for each end use. However, the results by end use allowed us to investigate whether savings uncertainty is related to measure type versus project size.

- **File review.** The file review involved extracting all project information relevant to savings estimation. This included:
    - Measure descriptions that detailed how the measure saves energy, affected systems and determinants of savings.
    - Baseline and efficient condition inputs to the M&V savings estimation tool, trend data, cutsheets and other design documents.
    - Reported savings values to compare against tracking data.
    - The final M&V savings estimation tool, and any other critical final documents used to document reported savings.
    - Invoices, receipts, and other data used to verify incremental measure costs.
    - Data and documentation relating to nonenergy benefits such as water use, wastewater, and operations and maintenance labor and materials.
    - Data used to determine nonelectric energy impacts.
    - Data to inform estimates of measure life.
- **Telephone/email discussion with project engineers.** The project engineers (BPA or utility) were another source of data. As needed, we contacted them by telephone or email to obtain information needed for the evaluation that was not found in the project files. These discussions informed practical strategies for minimally intrusive data collection from end users, and to clarify history and circumstances at the site. We discussed how operations may have changed because of the ongoing COVID-19 pandemic and if they were expected to impact savings calculations.
- **Telephone/email discussion with end users.** In most cases, it was necessary to obtain information from the end user via telephone or email. Discussions included operations staff and vendors to gather data baseline and post-installation conditions of affected buildings, systems and equipment. We also discussed how operations may have changed as a result of the ongoing COVID-19 pandemic. When necessary, these communications were used to plan site visits or remote data collection.
- **Site visits.** Based on the file review and discussions with internal engineers, we sometimes determined that more information was 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. For projects where site visits were challenging, we developed a more robust data collection survey that we administered via telephone and email with the appropriate end user and vendor staff. This included a greater reliance on file review findings, customer staff providing as-built plans and specifications, control system trend data and screen prints, and taking photos or videos and sending them to the evaluation team.
- **Affected system trend metering.** For custom projects, if there were insufficient trend data of critical systems to verify savings, additional metering data were collected. In most cases, this came from on-premises electric metering and other interval data correlated to savings such as air temperature or production levels. Interval premise data were collected from existing on-site instrumentation or from instruments installed by evaluators and on-site operations staff. Where on-site visits were not possible, we implemented a metering plan with the assistance of on-site staff. These plans leveraged existing metering and on-site staff with the skills necessary to install preconfigured data logging equipment.

- **Weather.** If weather data from the file review were not adequate, actual and/or typical meteorological year (TMY) weather data were acquired for the most appropriate National Oceanic and Atmospheric Administration (NOAA) weather station.
- **Cost effectiveness parameters.** To estimate measure cost effectiveness, we collected data for measure life, incremental costs, nonelectric energy use and nonenergy benefits. We relied on data found in file reviews; these only changed if there was compelling evidence found during evaluation. We did not reach out to end users solely about cost effectiveness parameters. Other cost effectiveness parameters including discount rates, administrative costs, and avoided energy costs used BPA-provided or default RTF values as necessary.

CUSTOM MEASURE ANALYSIS

We estimated savings for sampled custom measures as described below.

SELECT RELIABLE EVALUATION MODEL

Our starting point in estimating savings was a review of the M&V model. The first step was to determine whether the M&V model conformed to the BPA M&V protocols and RTF guidelines. It is important to note that determining compliance with a BPA M&V protocol was just the first step in reviewing an M&V model. The BPA M&V protocols and RTF guidelines 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 M&V protocol compliance and best practices were determined, we then examined the savings calculations in more detail to determine whether they provided the best practical estimate of savings.

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

Other areas germane to the model review included whether or not the M&V model addressed significant measure interactions, and whether or not it adequately established the proper baseline (current practice or precondition as defined in the RTF Guidelines[25] and the BPA M&V protocols). We considered interactions significant if it was likely that the interactive effect exceeds 10 percent of the measure savings.

If the M&V model was determined to be reliable, then we adopted it as the evaluation model, and improved its input data if necessary. If the M&V model was found to be unreliable, and thus not suitable to serve as the evaluation model, then we either enhanced or replaced it.

---

[25] Regional Technical Forum. 2018. *Regional Technical Forum Operative Guidelines for the Assessment of Energy Efficiency Measures*. https://nwcouncil.app.box.com/v/2018RTFOperativeGuidelines

Enhancement meant adding or replacing certain features, such as measure interactions, while replacement entailed 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 assumed that adequate measure information was available, and that in particular, data and analysis files were transparent and accessible. For example, if an M&V model was only available as a PDF file, then it would have been generally impossible for the evaluation team to assess the underlying algorithms and formulas for appropriateness and accuracy. While this did not occur, it would have become necessary to get further supporting details from knowledgeable staff, reconstruct the original model, or build an alternative one. Consequently, missing or inaccessible M&V models and supporting data could have led to substantial 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. We did encounter instances where the M&V model was not functional as supplied to account for changes in savings due to evaluation findings. To resolve those cases, we worked closely with BPA, ESI, and utilities to obtain the necessary information and to create an environment where the models could be executed.

ASSESS DETERMINANT RELIABILITY

Once the evaluation model—either the M&V model or a more reliable replacement—had been selected, we then considered each of the model inputs and determined what level of data collection was needed to support a sufficiently reliable savings estimate, as well as data collection needed for as-expected operating conditions. 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 developed 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 then found the corresponding values used in the evaluation model, assessed the data and/or documentation underlying those values, and determined whether we considered those values reliable. This involved some engineering judgment. To the extent that sampled measures involved similar systems, equipment and modeling techniques, we ensured that consistent judgments were applied.

For instance, we may have determined that hours of operation was 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 have concluded that the determinant value was reliable. However, if the metering only spanned two hours, we may have concluded it was unreliable, and therefore, additional on-site metering was necessary to develop a reliable determinant value. For each measure, we documented our rationale for establishing whether determinants were reliable or not, and how the reliability was to be improved if necessary. We compared across sampled measures to ensure consistency, as well as to identify overarching trends and issues.

We also verified determinant reliability in cases where energy efficiency performance was impacted by the COVID-19 pandemic. In the described case of a fan control measure, the facility may have added a shift due to increased production for pandemic-related essentials. We assessed determinant reliability for both the actual and as-expected scenarios.

For unreliable critical determinants, we assessed what level of data collection involving the end user was necessary to obtain reliability for that determinant. In order of cost and complexity, these levels were (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 determined the data collection approach for each sampled measure that established how data for each critical determinant were to be obtained. After our initial review, we prepared questions, data, and model requests for key personnel. This approach was incremental and iterative depending on availability of information and new findings, beginning with internal engineering personnel and included utility staff, vendors and customers. The iterative approach adhered to the contact protocols outlined in Appendix D: Utility & Customer Contact Protocols. Our general intent was 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. If data could not be produced, the best available information was used.

Certain measures required extended metering. Hypothetical examples included (a) a fruit processing facility with seasonal production schedules, or (b) a complex HVAC controls project that required separate summer and winter datasets to assess cooling and heating performance, respectively. Such instances were rare and were kept to a minimum because of the inconvenience to the end user, as well as the cost to the evaluation.

The data collection approach varied based on the types of data to be collected prior to and during the site visit. For example, a site visit may have involved interviews to find out about production seasons, coupled with collection of nameplate data and short-term metering. We structured our approach to collect data efficiently, with minimal impact on the end user. The approach also included unit sampling in situations where the measure consisted of many pieces of equipment.

RUN EVALUATION MODEL

If the M&V model was deemed appropriate to serve as the evaluation model, and the critical determinant values were deemed reliable, then this step was essentially 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 were corrected, and the proper values recorded for this evaluation. If a functioning savings model was not made available to the evaluation team, the team worked with vendors and program staff to make a model executable. If a more reliable model was needed, the evaluation team created one. All evaluation models were provided to the BPA evaluation team for review. Evaluated savings consisted of running the new or existing evaluation model with as-found reliable determinant values and as-found baseline operation conditions obtained through evaluation data collection.

To account for changes in savings resulting from the global COVID-19 pandemic, savings were calculated based on as-operated conditions described above, and as-expected conditions had COVID-19 not occurred. The intent was to also estimate what savings would have been in a non-pandemic year. In cases where savings were unchanged, the savings for both scenarios were identical. We searched for self-reporting of COVID-19 impacts during file review, and inquired further if necessary, during the phone/email interview steps. We investigated operational impacts such as temporary or permanent facility closure, changes in operating schedule, added or removed work shifts, increased outdoor air in HVAC schedules, and any other major operations change where a savings impact greater than 5 percent was expected. For example, a hypothetical plastics production facility with new injection molding machines produced 10,000 parts per day of PPE equipment. The savings was a function of the number of parts made and reduced energy consumption per part. The evaluated efficient case and baseline used 10,000 parts per day for the model. Before the pandemic, this facility planned to make 5,000 parts per day. We applied the same savings model with a baseline and efficient cases producing 5,000 parts per day as the expected savings model.

ESTIMATE SAVINGS USING ENGINEERING CALCULATIONS WITH VERIFICATION (ECWV)

We also used an ECwV protocol to estimate savings for each sampled measure. Our lead engineer for the site created a version of the site data that contained only the data needed for ECwV. In general, this eliminated long term post install trend data obtained from sub-metering and any conclusions reached by the analysis of such data. It was also occasionally necessary to substitute a different savings estimation model. Our team used the ECwV model to estimate savings and compare that to the best practical evaluation model results as well as the BPA ECwV estimate, where available. We determined the relative reliability of the two estimates. To account for changes in savings resulting from the global COVID-19 pandemic, ECwV savings were calculated as first year actual.

TREATMENT OF INTERACTIVE MEASURES

Savings achieved by one measure did affect the savings of another measure in the same project—for example, an HVAC upgrade and envelope improvements that affected the same spaces within a building. The change in envelope would have reduced heating and cooling losses. How much was saved by the HVAC upgrade could be substantially lower without the envelope change. Thus, the order in which savings were estimated could have made a difference for two measures (i.e., unique Technology/Activity/Practices at a single site). If the two improvements occurred as part of separate projects that were completed at different times, this was not an issue for this evaluation. Whichever measure we sampled, we accounted for the baseline conditions of the affected systems and equipment. If the HVAC measure was sampled for evaluation and the envelope improvements happened before the HVAC measure, our evaluation model would have captured the envelope characteristic as part of the baseline conditions.

TIME-BASED VALUE OF SAVINGS AND COST EFFECTIVENESS

There were a number of strategies for characterizing the time-based value of savings for the sample of measures. For this evaluation, we assigned load shapes to individual measures, as the current custom project calculator uses load shapes by sector. Using ProCost, we assigned each measure to one of the RTF savings shapes. We then calculated cost effectiveness and peak savings based on the generic calculator and project-specific ProCost analyses and reported on any differences.

STUDY AND DOMAIN ANALYSIS

Once data collection and analysis were completed for the sample, we compiled a workbook containing all of the individual site level quantitative outputs and qualitative findings about key drivers for deviations between evaluated savings and original savings estimates. These site-level results were then used to estimate the electric savings and cost-effectiveness for the domain and individual end uses within the domain using a ratio analysis. The current sample was stratified and designed to provide results for each of these end uses individually:[26]

- Compressed Air
- HVAC
- Motors/Drives
- Process Load
- Refrigeration
- Water Heating

For custom measures with ECwV protocol savings estimates, we also determined when the ECwV protocol provided a reliable estimate of savings.

FIRST-YEAR KWH SAVINGS

We estimated first-year savings for the domain and each end use based on the evaluation model results for the sample. Stratum-level realization rates were extrapolated to estimate savings for the remaining population within each stratum. Evaluated and estimated savings for the individual strata were then summed to estimate the overall domain results, enabling us to calculate an overall domain-level realization rate. Results were calculated and presented both with and without corrections for COVID-19-induced changes to operating behaviors.

RELIABLE SAVINGS FROM ECWV

As described above, we prepared two estimates of savings for each custom measure in the sample. We then compared the two savings estimates and examined the assumptions within the ECwV model and its inputs to assess the relative reliability of the ECwV estimate to answer these questions:

1. **Can the ECwV method be implemented to deliver both unbiased[27] and precise[28] estimates of savings?** In the domain analysis, we compared ECwV-evaluated results against as-observed evaluated results to determine whether there were systematic differences between evaluated results and ECwV results and enumerated the most influential drivers of those differences observed in the individual site results.
2. **Are there issues with how BPA applied ECwV protocols?** For projects that received a BPA-implemented ECwV estimate, we compared both ECwV evaluated results and non-ECwV evaluated results. This included applying the Protocol Selection Flowchart found on page 9 of the BPA *M&V Protocols Selection Guide*[29] and noting whether or not ECwV was applied according to the guide. Depending on the exact number of projects that we get with BPA ECwV estimates and the size of the difference between BPA-implemented ECwV and evaluation-implemented ECwV, we may or may not have observed statistically significant differences for the domain as a whole. In any case, we

---

[26] In BPA taxonomy, TAPs roll up into end-use groups. Therefore, where feasible, the evaluation attempted to roll up results into end uses for additional insight to BPA.
[27] Confirm total evaluated savings across sample for ECwV protocol are the same as for regular evaluation.
[28] Confirm individual measure-level ECwV and regular evaluation estimates correlate well.
[29] BPA. 2018. *Measurement & Verification (M&V) Protocol Selection Guide and Example M&V Plan*. https://www.bpa.gov/-/media/Aep/energy-efficiency/measurement-verification/1-bpa-mv-selection-guide.pdf (file will download automatically when clicking)

enumerated the most influential drivers of the differences observed in the individual site results to identify opportunities for improvement.

DOMAIN TIME-BASED VALUE OF SAVINGS AND COST-EFFECTIVENESS

In addition to first-year savings, the team extrapolated individual project impacts and cost-effectiveness results to the relevant stratum, end use, and total domain using the same ratio analysis framework. Because BPA provided all values necessary to calculate reported lifetime savings and reported cost-effectiveness, each of these values was used directly in the ratio estimation.

APPENDIX D: UTILITY & CUSTOMER CONTACT PROTOCOLS

This appendix describes the protocols the evaluation team used to contact utilities and end use customers while conducting the impact evaluation.

CONTACT PROTOCOLS

The evaluation team followed general end user and utility contact protocols for this evaluation related to contact with end users and utility representatives, including the following communication principles:

- Utilities are notified of their projects included in the evaluation before the start of evaluation activities and provided with clear information on samples, timelines and requirements. Utilities may reach out at this time to their customers to notify them of potential future contact by the evaluation team.
- 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 Contracting Officer's Technical Representative (COTR) and Account Executive (AE).
- Evaluation team provides at least one week of notice to utilities before any end user contact, including phone surveys and site visits.

UTILITY NOTIFICATION AND WEBINAR

Once the evaluation plan and sample were reviewed by the BPA evaluation team, the BPA evaluation Energy Efficiency Representative (EER) notified utilities via email that at least one project in their territory had been selected in the evaluation sample (either the primary or secondary sample). This initial email requested the primary utility contact for the evaluation and also included an invite to a webinar that included utility-specific information associated with the evaluation plan.

The evaluation team also provided detailed information to each utility about their sampled sites through a secured file transfer protocol (FTP); details were provided on custom project ID, utility-assigned ID, project name, facility name, address, completion date, sampled measure (TAP), invoice number, and whether the site was in the primary or secondary sample.

PROJECT DOCUMENTATION OR BILLING DATA REQUESTS

Because Option 2 utilities maintain the detailed custom project files, the evaluation team clarified the data request for all projects at the utility webinar and in written communication. Requested documents included:

- Final energy savings models
- Trend data used in models
- Cutsheets
- Invoices
- Other project related documentation

The evaluation team provided the utility with a timeline for file delivery (typically two weeks). The utility (or BPA, if requested by the utility) uploaded required files to a secure website.

To provide timely and actionable evaluation results, the team created an escalation protocol that was initiated when data collection efforts became delayed and posed an impact to the schedule. The protocol follows:

1. As noted above, initial data request emails were sent by the evaluation team to utility contact with a copy to the evaluation EER and evaluation lead.
2. If a utility requests more time, within the agreed-upon time limit, the utility EER and utility COTR are notified by the evaluation lead.
3. If a utility misses the deadline, then the evaluation EER, utility EER, COTR, and (potentially) the 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.

CONTACT OF INTERNAL PROJECT ENGINEERS

Following file review, the evaluation team contacted the internal (utility) project engineers to learn more about the project, on an ad hoc basis by the evaluation team. The discussion with the internal project engineer included:

- Answering questions regarding the project or files.
- Obtaining information needed for the evaluation that was not found in the project files.
- If end user contact was required, discussion was to inform the evaluation team on the history of the project and circumstances at the site and to identify the least intrusive approach for obtaining data needed by the evaluation.

PHONE SURVEYS, EMAILS, AND/OR SITE VISITS OF END USERS

Where the team utilized phone surveys, emails or site visits, the evaluation team lead engineer notified the utilities at least one week before any end user contact and provided them with a general description of information to be collected from the site. The phone survey and/or email collected relevant information and determined the necessity of site visits. The feasibility of on-site visits was at the discretion of the customer and the evaluation team. The evaluation team worked with BPA to develop materials to support any advance contact they wanted to make with end users, including email and phone call scripts. The evaluation team lead also provided a set of potential frequently asked questions to minimize any potential concerns by the end users.

Evaluation engineers followed reasonable safety and privacy requirements set by end users. This included safety training, personal protective equipment and health screenings. The site visit did not proceed until all reasonable end-user requirements for an on-site visit were met.