PACIFIC GAS & ELECTRIC ENERGY STAR RETAIL PRODUCTS PLATFORM (ESRPP)

PROGRAM PILOT EARLY EVALUATION

Final Report

January 18, 2019

CALMAC Study ID PGE0437.01

PRESENTED TO:

Doreen Caruth
Evaluation, Measurement and Verification
Pacific Gas & Electric
245 Market Street, N4Q
San Francisco, CA 94105

PRESENTED BY:

Brett Close, PhD
Managing Research Consultant
EMI Consulting
83 Columbia St. Suite 400
Seattle, WA 98104

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

APPENDICES

Executive Summary

PG&E ENERGY STAR Retail Products Platform (ESRPP) Program Pilot Early Evaluation

Introduction

This executive summary presents key findings from EMI Consulting's early evaluation of the PG&E ENERGY STAR Retail Product Platform (ESRPP) Program Pilot, covering the period 2016 through March 2018.

The PG&E ESRPP Program Pilot launched in March of 2016. It is a leading implementation of a larger national effort coordinated by program sponsors across the US. The Pilot aims to transform the market for select product categories of home appliances and consumer electronics towards higher efficiency by (1) influencing retailers to stock, sell, and demand more energy-efficient models in these product categories, and (2) collaborating with organizations to define more stringent specifications and standards.

Process evaluation objectives included: (1) Assessing and informing implementation of the program, (2) Validating key components of program theory, and (3) Providing data and information to aid the assessment of attribution, including information related to PG&E efforts to drive ENERGY STAR specifications.

Impact evaluation objectives included: (1) Measuring total program-qualified unit sales for participating retailers by product category/subcategory, (2) Measuring program-qualified share (PQS), or the percentage of total unit sales that are program-qualified by product category/subcategory, and (3) Computing program energy and demand savings.

Methods

- Program data review
- Interviews with PG&E staff and subcontractors (n=16 over two rounds)
- Interviews with external collaborators from collaborating organizations (n=8)
- Review of retailer implementation data
- Statistical modeling of retailer sales data
- Statistical modeling of in-store shelf assortment data

Program Period Covered:
March 2016 – March 2018

Summary of Findings

Process Findings: At this point in time, PG&E ESRPP Program Pilot processes are generally working well, though there are some areas where improvements could be made.

- All activities outlined by program theory have been successfully implemented, including the payment of incentives to participating retailers and the subsequent collection and tracking of sales data.
- PG&E has coordinated with the national ESRPP collaborative to select an optimal set of participating retailers and recruit new program sponsors, though it may be necessary to add more program sponsors going forward (in order to achieve greater program scale).
- The definition of product eligibility tiers is one area where program processes could be improved, as interviewees reported some challenges related to how ESRPP should define the appropriate levels.
- PG&E's participation in advocacy and outreach activities related to voluntary specifications (i.e., ENERGY STAR) is seen as impactful.

Impact Findings: Results show the PG&E ESRPP Program Pilot is leading to short-term and mid-term impacts as expected by program theory, though the results differ by product category and tier.

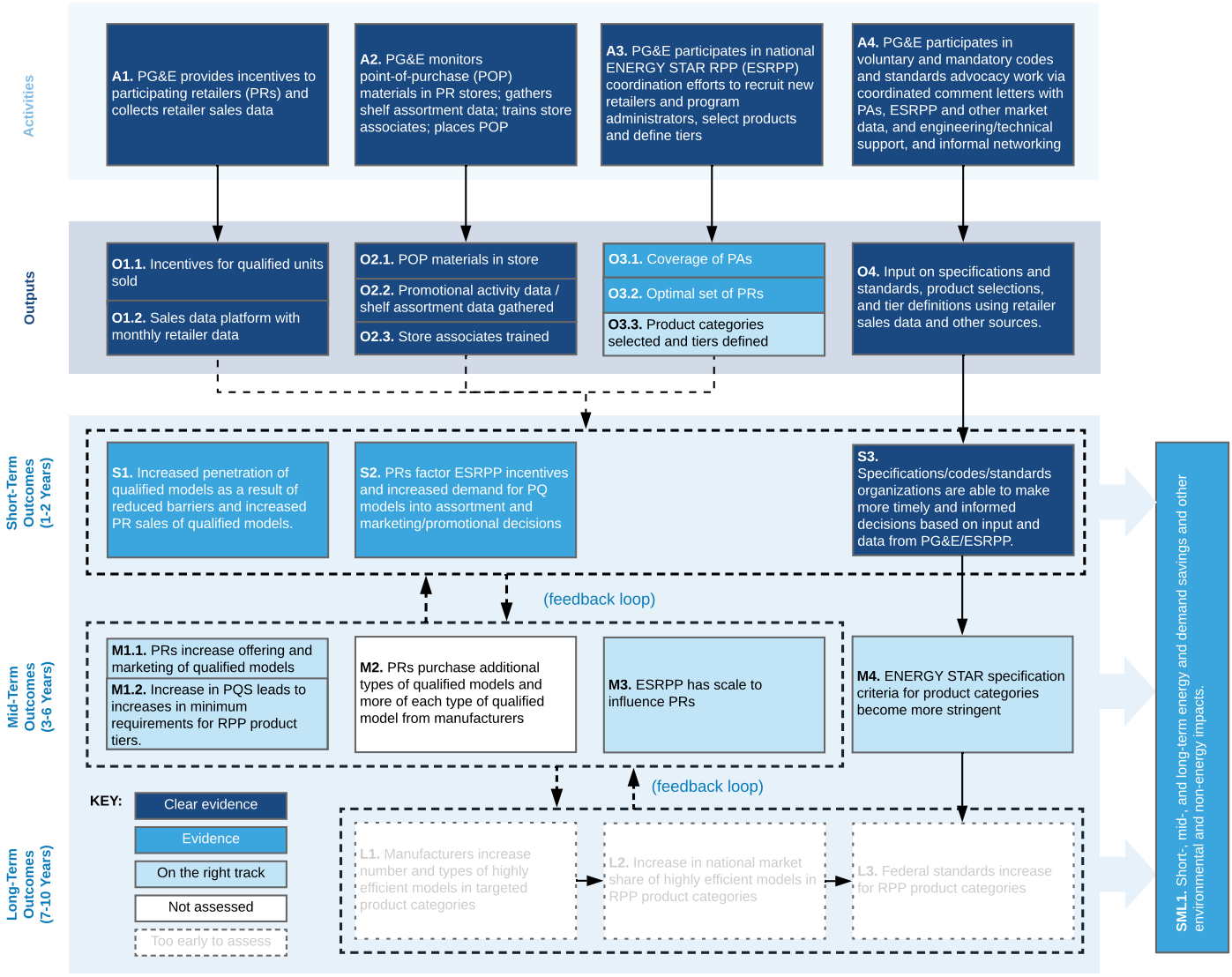
- Based on statistical modeling of retailer sales data, we observe short-term sales increases for 5/11 product tiers currently targeted by PG&E ESRPP.
- We also observe corresponding upward trends in program-qualified model assortment share on retailer shelves for five of seven categories.
- Interviews with national-level retail staff show that ESRPP incentives have some influence on retailer decision-making, and interviews with external collaborators show that ESRPP is facilitating the development of ENERGY STAR specifications.
- The national scope of ESRPP, coupled with the lack of availability of third-party market data outside of ESRPP, continues to present challenges to accurately determining program impacts.

PG&E ENERGY STAR Retail Products Platform (ESRPP) Program Pilot Early Evaluation

Logic Model Assessment

As part of the early evaluation, EMI Consulting worked with PG&E and its subcontractors to define a program logic model that clearly represents program activities and outputs, and maps them to desired short-, mid-, and long-term outcomes. This PG&E ESRPP logic model (shown below), highlights the importance of ESRPP contributions to the development of ENERGY STAR specifications and standards, and also highlights the complex, fluid nature of interactions between program outcomes.

Assessment of PG&E ESRPP Program Pilot Logic Model



Dashed lines denote multiple interactions between elements inside (i.e., arrows connecting all elements).

External Influences:
Broad economic conditions, market events, cost of energy, federal standards, ENERGY STAR, perceived need for conservation, and possible others.

Note: Factors can influence the program at all levels and time frames.

PG&E ENERGY STAR Retail Products Platform (ESRPP) Program Pilot Early Evaluation

Selected Impact Results

Results from the sales data regression modeling show increases for at least one tier for the following product categories: dryers, freezers, refrigerators, and soundbars. Results from the shelf survey analysis show preferential retailer treatment for all tiers except advanced freezers and advanced soundbars. Shelf survey analysis also shows upward trends in model assortment share for dryers, freezers, refrigerators, room air conditioners, and washers. In addition to stocking and assortment, program theory indicates that availability of incentives will lead retailers to provide qualified products with preferential treatment in their internal promotion decisions. Using model-level data collected from 403 individual store visits across 288 retail locations between January and August 2018, we calculated the percentage of models that were placed in a preferential location (anything other than simply in the aisle), the percentage of models that were currently discounted relative to the regular price, and the average discount amount among discounted products by product group and tier. Results indicate at least one type of preferential product treatment by retailers for 9 of 11 tiers currently incented by PG&E.

Summary of Impact Results

Product Category	Tier	Years Incented		Evidence of Preferential Treatment by Retailers?	Increase in Sales Above Baseline?	Increase in Availability of Program-Qualified Models on Shelves?
		PY 2016	PY 2017			
Air Cleaners	Basic	Yes	Yes	Yes	No	No
	Advanced	Yes	Yes	Yes	No	
Air Conditioners	Basic	Yes	Yes	Yes	Indeterminate	Yes
	Advanced	No	Yes	Yes	Too few sales	
Dryers	Basic	Yes	Yes	Yes	Yes	Yes
	Advanced	Yes	Yes	Yes	Yes	
Freezers	Basic	Yes	Yes	Yes	Yes	Yes
	Advanced	Yes	Yes	No	Too few sales	
Refrigerators	Basic	No	No	No	No	Yes
	Advanced	No	Yes	Yes	Yes	
Soundbars	Basic	Yes	No	Yes	No	No
	Advanced	Yes	Yes	No	Yes	
Washers	Basic	No	No	Yes	Yes	Yes
	Advanced	No	Yes	Yes	No	

Challenges and Limitations

The program pilot faced some major early hurdles related to data tracking and availability. Initially, challenges with the processing of retailer sales data made it difficult to perform thorough analysis of qualified models over time. This was more problematic for certain product categories where third-party data on the operating characteristics for individual models did not exist (air cleaners and soundbars). This processing has since become substantially more refined. Another early challenge was related to marketing plans that retailers were required to submit to the ESRPP program detailing their marketing plans for the upcoming program year. The idea is that these "retailer implementation plans" would enable evaluators to tie changes in sales back to specific activities undertaken by retailers and serve as evidence of attribution. However, in the course of early program development, it became clear that retailers were unable to provide this information in a format that would allow for the necessary analyses (likely because such marketing and promotional activities are not planned a year in advance). Instead, the program evaluation had to rely on in-store field data collected by a PG&E subcontractor to better understand what retailers were doing to promote and market program-qualified models.

Conclusions and Recommendations

The 2016-2018 PG&E ESRPP Program Pilot Evaluation research resulted in the following key conclusions and recommendations:

Conclusion 1: The PG&E ESRPP Program Pilot has implemented key activities necessary for the program to operate effectively, but impacts vary by product category. This reinforces the need for the program to have product category-specific strategies and goals that can be tracked and periodically reevaluated. It also suggests that not all product categories may be suitable to include in the ESRPP program.

> **Recommendation 1.1**: Continue to develop product category-specific strategies and targets that are tailored to each product. Additionally, for product categories where an increase in market share is not the primary objective, make sure that another objective has been identified and is clearly documented. For instance, there may be products where the primary objective is to help advance ENERGY STAR specifications. In these cases, there should be a specific need that ESRPP can address (for instance, by providing full-category sales data). To ensure that credit is given to PG&E, it is critical to document the impacts that these data have on subsequent developments for specifications, codes, or standards.
>
> **Recommendation 1.2**: Product categories for which we have not yet observed an increase in sales or assortment share should be closely monitored to ensure they are making reasonable progress toward the objective for that product category. For some product categories, the value of obtaining full category sales data from retailers may provide substantial benefit to PG&E efforts to advance specifications, codes, and standards. In these cases, there is an argument for keeping these product categories in the program, assuming that the relevant sales data can be used to advance voluntary or mandatory requirements (see Recommendation #1.1 above). It may be prudent to make downward adjustments to the incentive amounts for these product categories to reflect this strategy.

Conclusion 2: Analysis of sales data shows short-term increases in the sales share of program-qualified models for 5/11 product tiers, or 4/7 product categories currently targeted by PG&E ESRPP. At the same time, we observed preferential retailer promotional efforts and assortment increases for many of these same product categories. Collectively, this provides evidence that the core ESRPP program mechanism is working for these product categories/tiers. Our analysis indicates that the ESRPP intervention is linked to a statistically-significant increase in sales for dryers (basic and advanced), freezers (advanced), and soundbars (advanced). Additionally, we see a small but statistically-significant upward trend in the shelf assortment of program-qualified models on store shelves—a mid-term outcome which is expected to follow increases in program-qualified sales. Collectively these findings provide supporting evidence that, for some product categories, the core ESRPP intervention is having some effect.

Conclusion 3: National ESRPP program processes could be improved by adopting a simplified approach for defining tiers within a product category and, to the extent possible, aligning these tiers with ENERGY STAR requirements. An important feature of the ESRPP program design is the ability to "ratchet up" tier requirements as program-qualified share increases for these product categories. To date, the ESRPP collaborative has used a flexible method in which tier eligibility requirements are aligned annually with ENERGY STAR specifications except in cases where the market share for that product category is already high. In these cases, the tier requirements have been set to "ENERGY STAR + XX%" (where the precise percentage varies based on the current program-qualified market share). This is a necessary adjustment for the program to make. However, in some cases it has caused logistical difficulties for the program and for retailers because it becomes more difficult to determine which models actually qualify for each tier.

> **Recommendation 3.1:** In the future, PG&E should work with other program sponsors to explore simplifying the qualifying requirements used for the national ESRPP program and, to the extent possible, keeping these qualifying requirements aligned with ENERGY STAR definitions. For instance, ESRPP could choose to align qualifying requirements with ENERGY STAR Most Efficient (ESME) in categories where there is such designation. In categories that lack ESME, there may be value in working with the EPA to establish such a designation.

Conclusions and Recommendations (continued)

Conclusion 4: The full category sales data provided by participating retailers are a valuable tool, particularly for facilitating the development of specifications, codes, and standards. Interviews with external collaborators indicate that these data have already been used to facilitate the development of ENERGY STAR specifications. Further research has revealed that these data do not exist anywhere outside of the ESRPP efforts, making it an even more valuable resource.

> **Recommendation 4.1:** Given the long-term program goals of changing mandatory and voluntary specifications, PG&E should continue to work with regulatory bodies to provide data and analysis to accelerate the adoption of these rules.

Conclusion 5: PG&E's ESRPP program pilot is highly influential within the national ESRPP collaborative effort. Interviews with external collaborators provide evidence that PG&E is considered by other program sponsors and collaborating agencies to be one of two primary drivers of the national ESRPP effort, the other driver being the Northwest Energy Efficiency Alliance (NEEA). In particular, it appears that PG&E and NEEA are driving much of the codes and standards advocacy work.

Conclusion 6: As the PG&E ESRPP Program Pilot continues to operate moving forward, the current baseline approach (i.e., a pre/post model averaging baseline) will become less useful as the pre-period sales data become outdated. Therefore, it will become increasingly important to use a baseline approach that is able to account for new developments and external changes in the market.

> **Recommendation 6.1:** Moving forward, the PG&E ESRPP Program Pilot should adopt a baseline approach similar to that employed by NEEA to help understand and assess market transformation effects due to the ESRPP program. There are several benefits of using a baseline approach similar to that utilized by NEEA: (1) the approach has already been in use for some time, (2) it is transparent and flexible, and (3) using such an approach would facilitate evaluation consistency across two of the most important ESRPP program sponsors.

1. INTRODUCTION

Pacific Gas and Electric Company's (PG&E's) ENERGY STAR Retail Product Platform (ESRPP) Program Pilot aims to transform the market for select product categories of home appliances and consumer electronics towards higher efficiency by (1) influencing retailers to stock, sell, and demand more energy-efficient models in these product categories, and (2) collaborating with organizations to define more stringent specifications and standards.[1] The PG&E ESRPP Program Pilot is one of the first implementations of a larger national effort coordinated by program sponsors across the US.

This evaluation report provides the results of an early evaluation of the PG&E ESRPP Program Pilot by EMI Consulting, covering 2016 through March 2018. This introduction includes background information relevant to the PG&E ESRPP Program Pilot and the evaluation.

1.1 ESRPP PROGRAM HISTORY

Because plug loads represent a significant proportion of residential electricity consumption, reducing plug load energy consumption is a critical step on the path towards achieving California's residential Zero Net Energy (ZNE) goals. The 2012 ZNE Technical Feasibility Report stated that "…minimizing plug loads will be critical to meeting ZNE goals,"[2] and recommended that utilities "continue equipment efficiency incentive programs" and "aggressively promote equipment efficiency regulations at the state and federal level."[3]

The RPP concept was initially tested in a trial with a single participating retailer in 24 of its 41 stores located in the PG&E and SMUD service territories that took place

[1] California 2016-2019 Retail Products Platform Pilot Evaluation Plan. October 16, 2015.

[2] Arup, Davis Energy Group, Sun Light & Power, New Buildings Institute, Engineering 350, and Sustainable Design + Behavior. 2012. *The Technical Feasibility of Zero Net Energy Buildings in California.* Page 8. Developed on behalf of Pacific Gas & Electric Company. Retrieved from: http://www.energydataweb.com/cpucFiles/pdaDocs/904/California_ZNE_Technical_Feasibility_Report_Final.pdf

[3] Ibid. p. 51.

from November 2013 to December 2014.[4] The 2013-2014 RPP Trial incented six product categories, including: (1) air cleaners, (2) DVD/Blu-Ray players, (3) home theaters-in-a-box (HTIBs), (4) freezers, (5) refrigerators, and (6) room air conditioners.

1.2 CURRENT STATUS OF ESRPP

Starting in March 2016, the RPP Program became a national effort under the auspices of ENERGY STAR (henceforth referred to as ENERGY STAR RPP, or ESRPP). As of 2018, the participating retailers are Best Buy, The Home Depot, Sears/Kmart, Nationwide,[5] and Lowe's.

Utilities and energy efficiency organizations[6] ("program sponsors") across the US have partnered with each other to develop and implement ESRPP. Each participating program sponsor pays participating retailers per-unit incentives for every program-qualified unit sold during the program period. The program intent is to enlist additional program sponsors over time. As of March 2018, there were seven program sponsors.[7] With the addition of seven new program sponsors in mid-2018, the ESRPP program now covers roughly 18% of the total U.S. residential population.[8]

Within each product category, program-qualified models are divided into basic and advanced tiers based on efficiency levels set by the program sponsors. The models in the basic tiers meet or exceed the minimum ENERGY STAR specification; advanced tiers consist of more efficient models for which retailers receive higher

[4] A copy of the RPP Trial evaluation report is available here: https://www.etcc-ca.com/reports/pacific-gas-and-electric-company-retail-plug-load-portfolio-rpp-trial. The evaluation of the trial found roughly a 5% lift in qualified sales due to promotional activities tied to the intervention.

[5] Participating retailers are contractually obligated to provide sales data as part of the ESRPP program. Nationwide provides unit shipment data instead of unit sales data.

[6] Examples of non-utility program sponsors include Efficiency Vermont and the Northwest Energy Efficiency Alliance (NEEA).

[7] Program sponsors as of March 2018 included PG&E, the Northwest Energy Efficiency Alliance (NEEA), Sacramento Municipal Utility District (SMUD), Con Edison (New York), Focus on Energy (Wisconsin), Xcel Energy (Colorado and Minnesota), and Efficiency Vermont.

[8] New program sponsors onboarding in 2018 include four Maryland utilities (First Energy, Pepco, SMECO, and BGE) and two Connecticut utilities (Eversource and United Illuminating). Source: https://www.energystar.gov/sites/default/files/asset/document/ESRPP_1pager_08-29-18_508_0.pdf)

per-unit incentives.[9] The program theory holds that by increasing the sales of energy-efficient models over less efficient models, ESRPP will generate energy and demand savings for utility customers in the short-, mid-, and long-terms through participating retailers, while also transforming the overall market towards higher efficiency in the long-term.

For the 2016 national ESRPP Program cycle (March 2016 through March 2017),[10] Sponsors of the ESRPP Program incented five product categories: air cleaners, clothes dryers (both electric and gas), freezers, room air conditioners, and sound bars. For the 2017 program cycle (April 2017 through March 2018), two additional product categories were added to the program (clothes washers and refrigerators).[11]

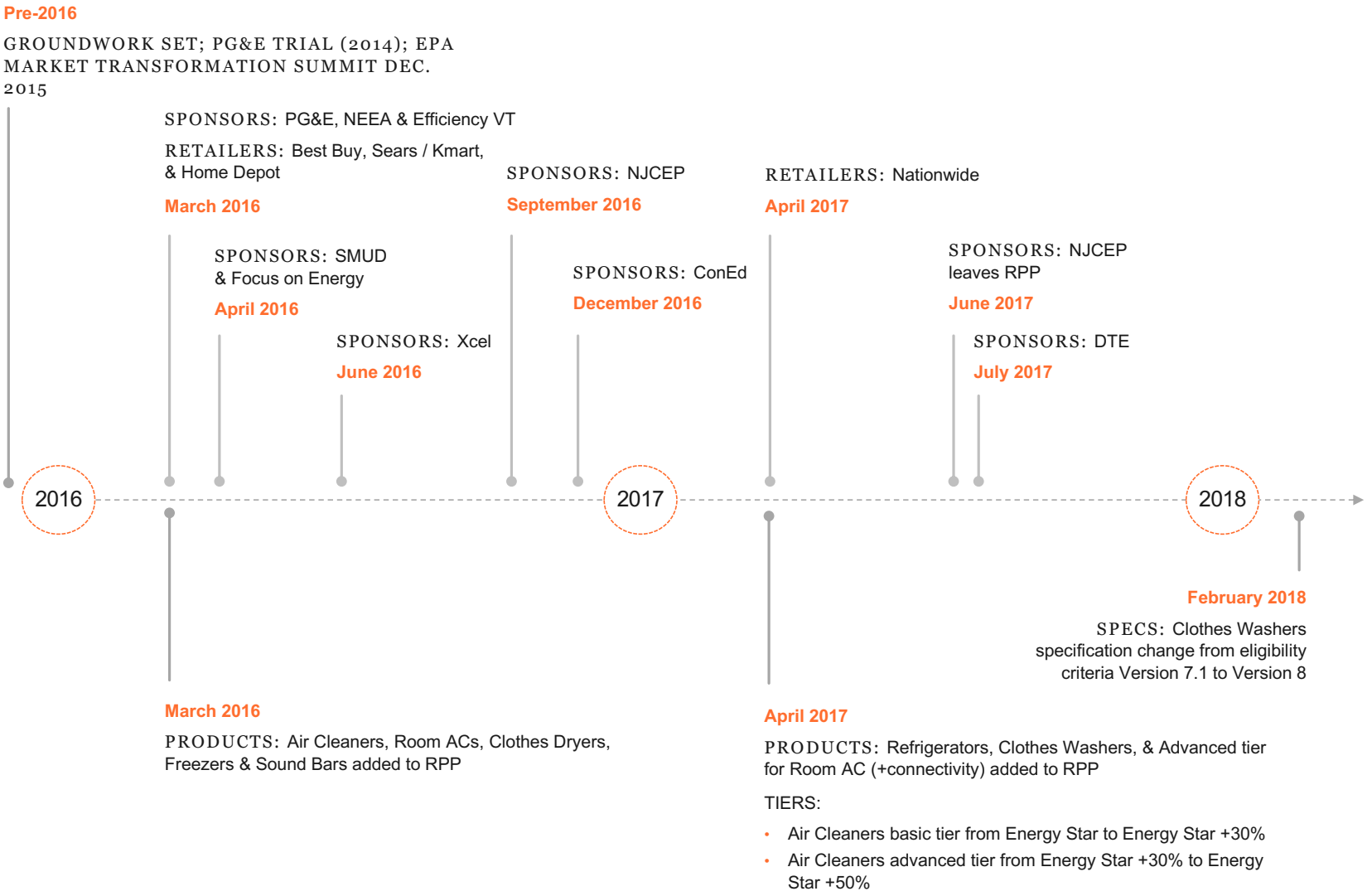
Figure 1-1 below provides a timeline of key ESRPP developments.

[9] The precise eligibility requirements for basic and advanced tiers vary by product category and are based on each product's unique set of efficiency metrics (e.g., clean air delivery rate for air cleaners).

[10] The national program cycle runs from April through March, though the first-year started in March 2016; new retailers or program sponsors can join at any time.

[11] Dehumidifiers were added as a national-level product category in April 2018 but are not included in the scope of this evaluation.

Figure 1-1. National ESRPP Timeline, March 2016-March 2018



PG&E IMPLEMENTATION OF ESRPP

As one of the first and largest ESRPP program sponsors, PG&E manages a portfolio of seven product categories. Each product category is divided into a basic and advanced tier, though PG&E does not incent all tiers. As shown below in Table 1-1, each tier has an associated incentive value. From the 2016-2017 program cycle to the 2017-2018 program cycle, the tier requirements changed for air cleaners, dryers (advanced tier), refrigerators (advanced tier), room ACs (advanced tier), and washers (advanced tier). PG&E additionally adjusted its incentive amounts for dryers, refrigerators (advanced tier), and washers (advanced tier).

Table 1-1. PG&E ESRPP Product Categories, Tiers, and Incentives, 2016-2018

Product	RPP Tier	2016		2017	
		Spec	Incentive	Spec	Incentive
Air Cleaners	Basic	ES v1.2	$20	**ES v1.2 +30%**	$20
	Advanced	ES v1.2 +30%	$30	**ES v1.2 +50%**	$30
Clothes Dryers	Basic	ES v1.0	$50	ES v1.0	**$30**
	Advanced	ES ET Award (electric only)	$250	**ESME 2017**	$250
Freezers	Basic	ES v5	$20	ES v5	$20
	Advanced	ES v5 +5%	$50	ES v5 +5%	$50
Refrigerators	Basic	-	-	ES v5	$0
	Advanced	-	-	**ESME 2017**	**$20**
Room Air Conditioners	Basic	ES v3.1	$20	ES v4	$20
	Advanced	-	-	**ES v4 + conn.**	$20
Soundbars	Basic	ES v3 +15%	$10	ES v3 +15%	$0
	Advanced	ES v3 +50%	$20	ES v3 +50%	$20
Washers	Basic	-	-	ESME 2017	$0
	Advanced	-	-	**ESME 2017 +5%**	**$20**

Note: Tiers or incentives that changed in the second year of program operations are in bold.

1.3 EVALUATION OBJECTIVES

Because the PG&E ESRPP Program Pilot concept is one of the first programs of its type aimed at longer-term market transformation in the State of California,[12] the evaluation has assessed various program processes in addition to identifying and measuring performance and market transformation indicators.

The objectives of the **process evaluation** of the PG&E ESRPP Program Pilot include:

- Assessing and informing the implementation of the program
- Validating key components of the program theory
- Providing data and information to aid the assessment of attribution, including the degree to which PG&E collaborates with other organizations to define more stringent specifications and standards

Since the primary performance objectives of the PG&E ESRPP Program Pilot are to increase sales of energy-efficient products that will, in turn, affect reductions in energy consumption for targeted product categories, the PG&E ESRPP Program Pilot evaluation includes an **impact evaluation** to inform savings claims,[13] which will be aimed at accurately:

- Measuring total program-qualified unit sales for participating retailers by product category/subcategory
- Measuring program-qualified share (PQS), or the percentage of total unit sales that are program-qualified by product category/subcategory
- Computing gross program energy and demand savings

To the extent possible, the evaluation of the ESRPP Program Pilot complies with the California Energy Efficiency Evaluation Protocols: Technical, Methodological, and

[12] The PG&E ESRPP Program Pilot is a larger-scale version of the PG&E RPP Phase I Trial, which took place between 2013-2014. The Phase I Trial followed the Business and Consumer Electronics (BCE) program, which took place in 2010-2012.

[13] Note that this is not an impact evaluation per se, as the CPUC-ED will lead any official impact evaluation of the RPP Program. The impact evaluation as discussed herein is intended to serve as an early M&V effort aimed at assessing and informing savings estimation and attribution prior to any actual ED-led impact evaluations.

Reporting Requirements for Evaluation Professionals[14] and The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users.[15] However, because of the diversity of evaluation objectives that exist for the ESRPP Program, no single methodology is suitable for assessing all objectives. Some objectives are more qualitative in nature and involve assessing and evaluating operational activities and processes to ensure that the program is being implemented as planned and functioning as expected. Other objectives are more quantitative in nature and involve defining, measuring, and analyzing specific indicators of program progress, attribution, and/or success. Also, since this is a novel program concept aimed at market transformation, an additional objective of the evaluation includes assessing the array of potential approaches to evaluating the program to identify the most effective, informative, and feasible approaches to apply, should the program be further scaled up in ensuing years.

Because the California Public Utilities Commission's Energy Division (CPUC-ED) is responsible for conducting *ex post* impact evaluations in California, this evaluation of the PG&E ESRPP Program Pilot should be viewed as an Early EM&V effort, as permitted for pilot programs in California. This evaluation also incorporates lessons learned from leading the evaluation of the initial 2013-2014 RPP Program Trial,[16] as well as lessons learned from the first two years of pilot program operations. As such, it should also be viewed as a second-phase developmental evaluation.[17] Additionally, the results of this evaluation may be used to assist other Program Administrators and ENERGY STAR in the development and implementation of the multi-region evaluation efforts.

[14] TecMarket Works Team, 2005. California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals. Prepared for the California Public Utilities Commission.

[15] Yarbrough, D. B., L. M. Shulha, R. K. Hopson and F. A. Caruthers. 2011.The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users. Los Angeles, CA: Sage Publications.

[16] Malinick, T. and Ridge, R. 2015. Pacific Gas and Electric Company Retail Plug-Load Portfolio (RPP) Trial: Evaluation Report. April 24, 2015.

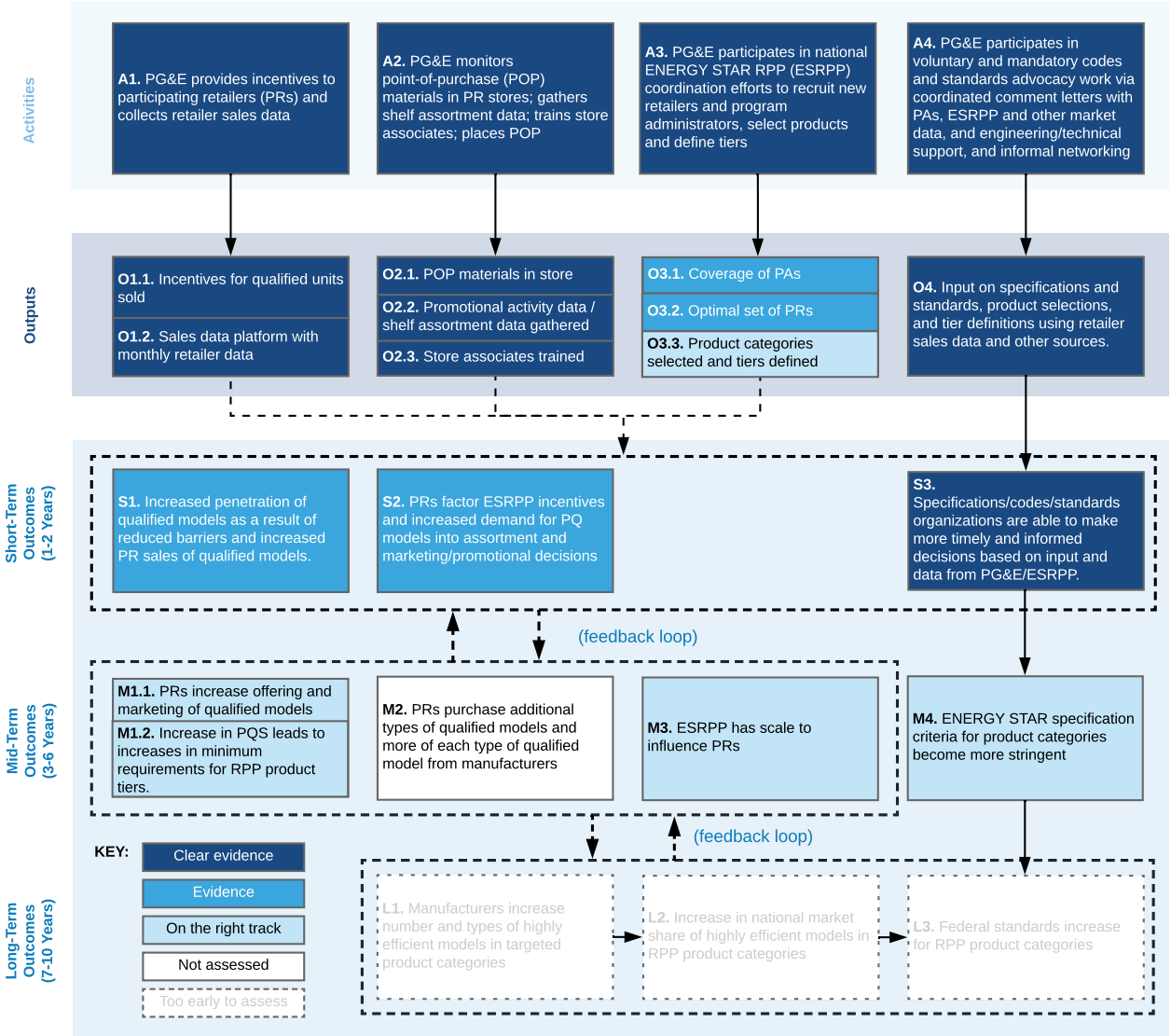
[17] Patton, M.Q. 2010. Developmental Evaluation. Applying Complexity Concepts to Enhance Innovation and Use. New York, NY: Guilford Press.

1.4 KEY FINDINGS

In this section we provide key findings from the process and impact evaluations of the PG&E ESRPP Program Pilot. An overall summary graphic is provided below as Figure 1-2.[18] We then discuss key process findings and key impact findings.

[18] This graphic represents a set of proposed revisions to the existing program logic model. More information on these revisions is provided in Appendix F.

Figure 1-2. Graphical Summary of PG&E ESRPP Program Pilot Evaluation Results



KEY PROCESS FINDINGS

The PG&E ESRPP program pilot faced some major early hurdles related to data tracking and availability. Initially, challenges with the processing of retailer sales data made it difficult to perform thorough analysis of qualified models over time. This was more problematic for certain product categories where third-party data on the operating characteristics for individual models did not exist (air cleaners and sound bars). This processing has since become substantially more refined. Another early challenge was related to marketing plans that retailers were required to submit to the ESRPP program detailing their marketing plans for the upcoming

program year. The idea was that these "retailer implementation plans" would enable evaluators to tie changes in sales back to specific activities undertaken by retailers and serve as evidence of attribution. However, in the course of early program development, it became clear that retailers were unable to provide this information in a format that would allow for the necessary analyses (likely because such marketing and promotional activities are not planned a year in advance). Instead, the program evaluation had to rely on in-store field data collected by a PG&E subcontractor to better understand what retailers were doing to promote and market program-qualified models.

At this point in time, PG&E ESRPP Program Pilot processes are generally working well, as reported by staff and external collaborator interviewees. As shown in the graphical PG&E ESRPP logic model above in Figure 1-2, all activities and outputs are occurring as intended. The incentive payments to participating retailers and the subsequent collection of sales data—collectively the core "engine" of ESRPP—are occurring (logic model elements A1 and O1). Similarly, the collection of in-store field data and placement of ESRPP signage by PG&E has been successfully implemented (logic model elements A2 and O2). PG&E has successfully coordinated with the national ESRPP collaborative to recruit new program sponsors and select an optimal set of participating retailers, though some interviewees believe that the addition of more program administrators is necessary going forward (in order to achieve greater program scale, which would bring more leverage, and thereby greater transformation). The definition of product eligibility tiers (logic model element O3.3) is one area where program processes could be improved, as interviewees reported some challenges related to how ESRPP should define the appropriate levels. Lastly, there is clear evidence that PG&E's participation in activities related to voluntary specifications, codes, and standards (logic model elements A4 and O4) is working well and leading to the intended outputs.

KEY IMPACT FINDINGS

Early evaluation results provide evidence that the PG&E ESRPP program is leading to short-term and mid-term impacts as expected by program theory, though the results differ by product category. Based on statistical modeling of retailer sales data, we observe short-term sales increases (logic model element S1) for four of the seven product categories currently targeted by ESRPP, and corresponding upward trends in program-qualified model assortment share on retailer shelves for five categories (logic model element M1.1). At the same time, interviews with national-level retail staff show that ESRPP incentives have some influence on retailer decision-making (logic model element S2), and interviews with external collaborators show that ESRPP is facilitating the development of ENERGY STAR specifications (logic model element S3). At this point in time, it is premature to assess the long-term outcomes included in the logic model.

There are also a number of shortcomings identified in the program pilot's ability to achieve its desired short-term and mid-term outcomes.

- In looking at short-term outcomes, we do not see statistically-significant increases in sales across *all* product categories and tiers receiving incentives (logic model element S1).[19]
- Secondly, although retailers do indicate that ESRPP incentives have factored into their decision-making (logic model element S2), it remains difficult to understand exactly how the ESRPP incentives are considered by retailers relative to other considerations (e.g., manufacturers competing for shelf space).
- Lastly, at this stage, much of PG&E ESRPP's advocacy efforts have been aimed at ENERGY STAR, with limited activities aimed at other standards-setting bodies (logic model element S3).

In looking at mid-term outcomes, we see several additional shortcomings.

- While our analysis of shelf survey data shows increases in model assortment share for 5 of 7 product categories, an additional two product categories are either flat (air cleaners and soundbars) or have a decreasing trend in model assortment share. The reasons for this are unclear. PG&E has identified that the primary objective for these categories is to facilitate the advancement of ENERGY STAR specifications through the provision of market data, making the lack of increasing market share for these categories less critical.
- In the mid-term, increasing market share is designed to lead to a "ratcheting up" of program requirements. In many cases this requires setting an efficiency level based on ENERGY STAR but does not map directly to an existing designation such as ENERGY STAR Most Efficient. This new level instead takes the form of "ENERGY STAR + XX%" and makes it difficult for retailers (and potentially customers) to easily understand which models are program-qualified without looking at the official qualified products list.

In the next chapter, we outline the methods used as part of this evaluation.

[19] As noted elsewhere in this report, for some product categories/tiers, increasing sales of program-qualified models was not the primary objective.

2. METHODS

This chapter includes a summary of methods used to complete the evaluation research. Research activities included:

- A review of PG&E ESRPP Program documentation
- Interviews with PG&E ESRPP program staff and external collaborators
- A review of interviews conducted with national-level retail staff [20]
- A review of retailer-provided implementation plans
- Regression analysis using sales data provided by participating retailers
- Shelf survey analysis using field data collected by PG&E

Table 2-1 below indicates the timeline showing when each of these research activities was conducted.

[20] These interviews were conducted by Cadmus as part of the national ESRPP program efforts.

Table 2-1. Timeline of Evaluation Research Activities

Research Activity		Timeframe	General or Product Category-Specific?
Review of PG&E Program Documentation		EMI Consulting has reviewed program strategy documents at several points throughout 2018.	Mainly general. Some documents did have category-specific information.
In-Depth Interviews with PG&E ESRPP Staff and External Collaborators	Interviews with PG&E ESRPP Staff	EMI Consulting initially conducted a round of interviews with PG&E ESRPP staff in late 2016 (n=8). EMI Consulting conducted a second round of interviews with PG&E ESRPP staff in mid-2018 (n=8). The second round of interviews included some but not all of the staff from the first round.	General
	Interviews with External Collaborators	EMI Consulting conducted eight in-depth interviews with "external collaborators" in mid-2018. These collaborators mainly represented entities involved with specification development, as well as other ESRPP program sponsors. These interviews also included two manufacturing suppliers.	General
Interviews with National-Level Retailer Staff		Year 1 Interviews were conducted by Cadmus in late 2016 and early 2018. The results of these interviews were then provided to EMI Consulting.	Mainly general
Review of Retailer Implementation Plans		Plans were provided by retailers at the beginning of PY 2016 and PY 2017. EMI Consulting reviewed these plans at that time.	Plans were intended to be product category-specific but the documents provided only allowed a more general analysis.
Regression Analysis of Retailer-Provided Sales Data		EMI Consulting conducted a regression analysis of retailer-provided sales data in late 2018.	Product-category specific
Analysis of In-Store Field Data Collected by PG&E		EMI Consulting conducted an analysis of in-store field data in late 2018.	Product-category specific

In Table 2-2 below, we map the research activities to the evaluation objectives of this study.

Table 2-2. Mapping of Research Activities to Evaluation Objective

Objective	Evaluation Activity					
	Program documentation review	Interviews with staff and external collaborators	Review of national retailer interviews	Review of retailer implementation plans	Sales data analysis	Shelf survey analysis
Assess and inform the implementation of the program	●	●	●	●	●	●
Validate key components of the program theory	●	●	●	●	●	●
Provide data and information to aid the assessment of attribution	●	●	●	●	●	●
Measure total program-qualified unit sales for participating retailers by product category/subcategory					●	
Measure program-qualified share by product category/subcategory					Program-qualified sales share	Program-qualified model assortment (shelf) share
Compute program energy and demand savings					●	

2.1 PROGRAM DATA REVIEW

As part of the ESRPP Program Pilot, PG&E engages Energy Solutions to formulate and implement strategies for each product category, with the goal of maximizing program influence and understanding how efficiency requirements should be set (i.e., at what level). The evaluation team reviewed documentation of these product strategies (called "Product Strategy decks") provided by Energy Solutions to better understand how the PG&E ESRPP Program Pilot is designed to influence product categories.

PG&E also engages with ICF International, which serves as the retail data aggregator for the national ESRPP effort. ICF International receives retailer data from retailers and processes this data to determine which model sales are qualified vs. non-qualified, and to assign them to the appropriate tier, bin, and any other subcategories as needed. ICF International then makes this processed data available to program sponsors and evaluators via an online data portal. Retailer sales data is subject to a number of restrictions designed to ensure a degree of anonymity. For example, retailer names are not attached to sales of non-qualifying models.

2.2 INTERVIEWS WITH PROGRAM STAFF AND EXTERNAL COLLABORATORS

As part of the implementation of the ESRPP Program, Program Administrator (PA) staff and the implementation team are expected to network with different organizations, such as:

- Internal PG&E Codes & Standards program staff
- Staff from other ESRPP PAs
- EPA/ENERGY STAR staff, and
- Staff from agencies involved with setting codes and standards

As part of the PG&E ESRPP Program Evaluation, EMI Consulting conducted eight PG&E ESRPP staff interviews and eight external collaborator interviews. The program staff interviewees included four PG&E program staff and four subcontractor staff. The external collaborator interviews included two other ESRPP Program Administrators (i.e., not PG&E); three staff members of government agencies working with specifications, codes, and standards; and two upstream suppliers that provide components to manufacturers. Interviews were conducted over the phone and typically lasted between 45 minutes and 1 hour.

The specific research topics/questions for this task included the following:

- What does communication look like within the PG&E ESRPP Program Pilot?
- What are the lessons learned from the development and implementation of the pilot? What have been the greatest successes? What have been the greatest challenges?
- Are the program processes appropriate and scalable?
- What is the level of effort PG&E staff is putting towards program activities? According to external collaborators, how influential is the program pilot?
- What is the level of satisfaction with PG&E ESRPP?
- What are recommendations for change and improvements to the program moving forward?

Based partially on the results of these interviews, as well as on a review of program operations and discussions with program staff, EMI Consulting designed a revised logic model to better represent program operations moving forward.

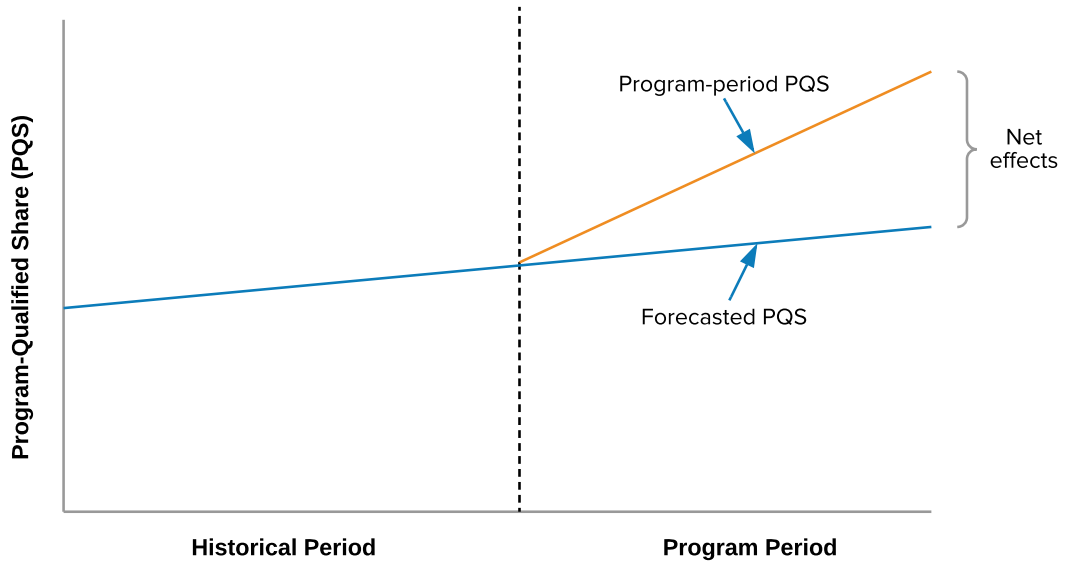
2.3 REVIEW OF RETAILER INTERVIEWS AND RETAILER-PROVIDED IMPLEMENTATION DATA

As part of the national ESRPP evaluation effort, Cadmus conducted 60-minute interviews in late 2016 with retail merchants (responsible for retailer purchasing decisions), marketing staff, and sustainability specialists from Best Buy, Sears/Kmart, and The Home Depot. The interviews addressed all product categories targeted at that time: air cleaners, room air conditioners, sound bars, clothes dryers, and freezers. Cadmus performed follow-up interviews with national retailers in early 2018, covering many of the same topics addressed in the first round of interviews. EMI Consulting reviewed the results of both sets of interviews as part of this evaluation. Additionally, EMI Consulting reviewed retailer implementation plans (RIPs) provided by participating retailers.

2.4 SALES DATA REGRESSION ANALYSIS

EMI Consulting estimated changes in unit sales for ESRPP product groups through the use of a pre/post model-averaging baseline comparison. This analysis involved creating a statistical model of sales in the pre-program period, using this statistical model to predict sales during the program period, and then comparing the predictions to the observed sales. This is depicted graphically below in Figure 2-1.

Figure 2-1. Simplified Depiction of Pre/Post Modeled Baseline Approach



EMI Consulting adjusted the data for seasonal variation using the observed seasonal patterns over time to smooth out the variation and adjust for different likely sales volumes during different parts of the year. Then, for each product group and classification tier, we developed three statistical models of baseline sales behavior that incorporate different assumptions about how the program affects qualified product sales and how the baseline sales behavior changes. Each of these models allow for "naturally occurring" pre-program trends in sales or market share and is evaluated on the pre-program sales data. The three statistical models are detailed below:

- **Sales Model**: This model uses monthly sales values, assuming that the effect of the program is to increase the sales of program-qualified products. This model explicitly allows qualified and non-qualified sales to vary separately.

- **Market Share Model**: This model uses monthly penetration rates, assuming that the effect of the program is to increase the market share. This model combines the qualified and non-qualified model sales and relies on changes in the ratio over time.

- **Probit Model**: This model uses a transformation of the market share used in the Market Share Model. It assumes that the effect of the program would have a smaller absolute impact on the market share if it is extreme (either very small or very large), but a larger absolute impact if the market share is modest.

For each product group, we combined the three models into an 'Averaging Model' to predict sales levels during the program period.

- **Averaging Model**: This model combines the three models above, based on how well the models predicted actual sales for the pre-program period, to develop a predicted sales value for each program group. This allows the model to incorporate the different assumptions in the three previous models to be combined in a way that fits the data best.

The weights for combining the models are selected using a numerical optimization routine to minimize a statistical measure called the leave-one-out cross validation criterion, which is a measure of how well the model fits each observation if it does not include that observation in the data it uses.[21] We estimated increases in qualified product sales as the difference between the observed sales and the predicted sales. If observed sales were larger than predicted sales, then that constituted an increase in the qualified product sales level during the program period. We then determined if the predicted increases were statistically significant with at least 90% confidence. Because the increase is relative to the baseline market behavior before the program started, we consider the increase attributable to the program.

ESTIMATING PER-UNIT ENERGY SAVINGS

To calculate energy savings, EMI Consulting relied on the per-model savings estimates in the ICF data portal.[22] For each product category we calculated the weighted average of the electric unit energy savings (kWh), electric unit demand reduction (kW), and gas unit energy savings (therms). We then took the weighted average savings for all qualified models sold in a given product category during the program period and multiplied this average value by the sales increase to determine overall savings for that category. The total energy savings and demand reductions are the product of the sales increase and the unit energy savings or demand reduction. We treated unit energy savings and unit demand reduction values as fixed, rather than uncertain, so that the uncertainty in the final savings

[21] This method of model averaging is known as Hansen-Racine Jackknife Model Averaging: Hansen, Bruce E. and Jeffrey S. Racine. "Jackknife model averaging." *Journal of Econometrics*, 167 (2012) pp. 38-46.

[22] EMI Consulting performed a comprehensive check on these values in mid-2017 to ensure that the values in the portal matched the values in the appropriate PG&E workpaper.

estimates is based on the uncertainty in the program-induced sales increase, and not the uncertainty in the energy savings values.

CHALLENGES AND LIMITATIONS

The primary challenge of estimating sales increase is lack of data in the pre-program period. Because product groups have between 12 and 25 months of pre-program sales data, adjustments for seasonality and pre-existing trends in qualified product sales require making assumptions about the underlying seasonality and trend behavior, albeit based on the observable patterns in the data. For example, if there are two observations from the month of June and those observations are high, we assume that those are June seasonal effects and not random fluctuations or due to some other cause. And because no comparison group data are available, the only comparison we can make is based on using the pre-program period to inform what we think would happen to sales in the absence of the program intervention. The approaches taken by EMI Consulting are described briefly below and in more detail in Appendix A.

For any program like ESRPP, data quality is an ongoing challenge. While data quality improved markedly over the course of the first two years of program operations, earlier versions of the data portal did not always classify models consistently over time (as might be needed by evaluators). EMI Consulting conducted an initial quality assurance/quality control (QA/QC) review of the data provided via the data portal and posed questions to the data aggregator, where applicable. Following these discussions, EMI Consulting operated under the assumption that the sales data downloaded from the data portal were correct in terms of sales numbers, program-qualified status, and energy savings.

A broader challenge likely to affect future ESRPP evaluations is that the program design is based only on in-store sales and does not consider the effects of online sales of models in program-eligible product categories. Online sales are becoming increasingly important as more customers make purchases online (at least for some product categories, such as soundbars). It will be important for staff to understand how the program is affected by this channel.

Lastly, it is important to note that the modeling performed as part of this evaluation does not factor in efforts by PG&E and other IOUs over the past few decades to advance energy efficiency through a number of different programs. It is difficult to

estimate the precise impacts that these many programs had on efficiency levels for ESRPP products, though it is possibly quite large.[23]

2.5 SHELF SURVEY ANALYSIS

Through in-store field visits at participating retail locations, PG&E has gathered data on product assortment. Data collected during these in-store visits include information on which models were stocked on store shelves at each retail location on a monthly basis. In our analysis of shelf-survey data, we identified the number of unique models (within a product category) that were program-qualified (by tier) and the number that were non-qualified. This activity allowed us to look at changes in the proportion of distinct models in retailer assortments that are program-qualified over time.

2.6 SYNTHESIS

Throughout this research, the evaluation team relies on a theory-driven evaluation approach[24] to bring together the results of the process evaluation and the impact evaluation. This approach involves operationalizing the key performance indicators associated with key causal linkages in the logic model (i.e., converting the performance indicators into quantifiable and measurable indicators). If the predicted steps between program activities, outputs, and outcomes can be confirmed in implementation, then this matching of the theory to observed outcomes lends a strong argument for causality. As presented throughout this report, the logic model and underlying program theory have guided the evaluation in order to understand whether the PG&E ESRPP Program Pilot is functioning as intended. While some of these analyses will support reliable conclusions about short-term activities, outputs, and short-term outcomes, an assessment of other mid- and long-term outcomes must be supported through comparisons with the results of future evaluations.

[23] For example, see: "Energy Efficiency Portfolio Report." California Public Utilities Commission. May 2018. Available: http://www.cpuc.ca.gov/uploadedFiles/CPUCWebsite/Content/About_Us/Organization/Divisions/Office_of_Governmental_Affairs/Legislation/2018/13-15%20Energy%20Efficiency%20Report_Final.pdf

[24] For more information on theory-driven evaluation see: Weiss, C.H. 1997. "Theory-based Evaluation: Past, Present and Future." In: D.J. Rog & D. Fournier (Eds.), *Progress and Future Directions in Evaluation: Perspectives on Theory, Practice and Methods* (pp. 41-55).

3. PG&E ESRPP PROGRAM THEORY

In this chapter we first provide a brief primer on the theory underlying the PG&E ESRPP Program Pilot, including a discussion on how this theory has changed over time. We also provide a detailed explanation of the main levers that ESRPP uses to impact the market.

3.1 PG&E ESRPP PROGRAM THEORY

There are two important aspects of the ESRPP program design:

- **ESRPP is national in scope.** Program sponsors across the US strive for a consistent portfolio of qualifying products in home appliance and consumer electronics categories and coordinate implementation across regions.
- **ESRPP uses a mid-stream delivery mechanism as leverage to influence manufacturers and accelerate the development of specifications, codes, and standards**. Per-unit incentives are paid to participating retailers, with the goal of influencing manufacturers. Additional activities are aimed at influencing development of specifications, codes, and standards.

ESRPP program theory is based on the notion that collective incentives create scale to motivate retailers to assort and sell more qualified models, eventually leading to more orders of energy-efficient models to manufacturers. Specifically, this series of reactions will lead to:

- **Energy and demand savings** for utility customers in the short-, mid-, and long-terms.
- **Market transformation** that grows the customer energy savings opportunity in the long-term as increased sales impacts manufacturing and higher market penetration creates more stringent ENERGY STAR specifications and federal codes and standards.

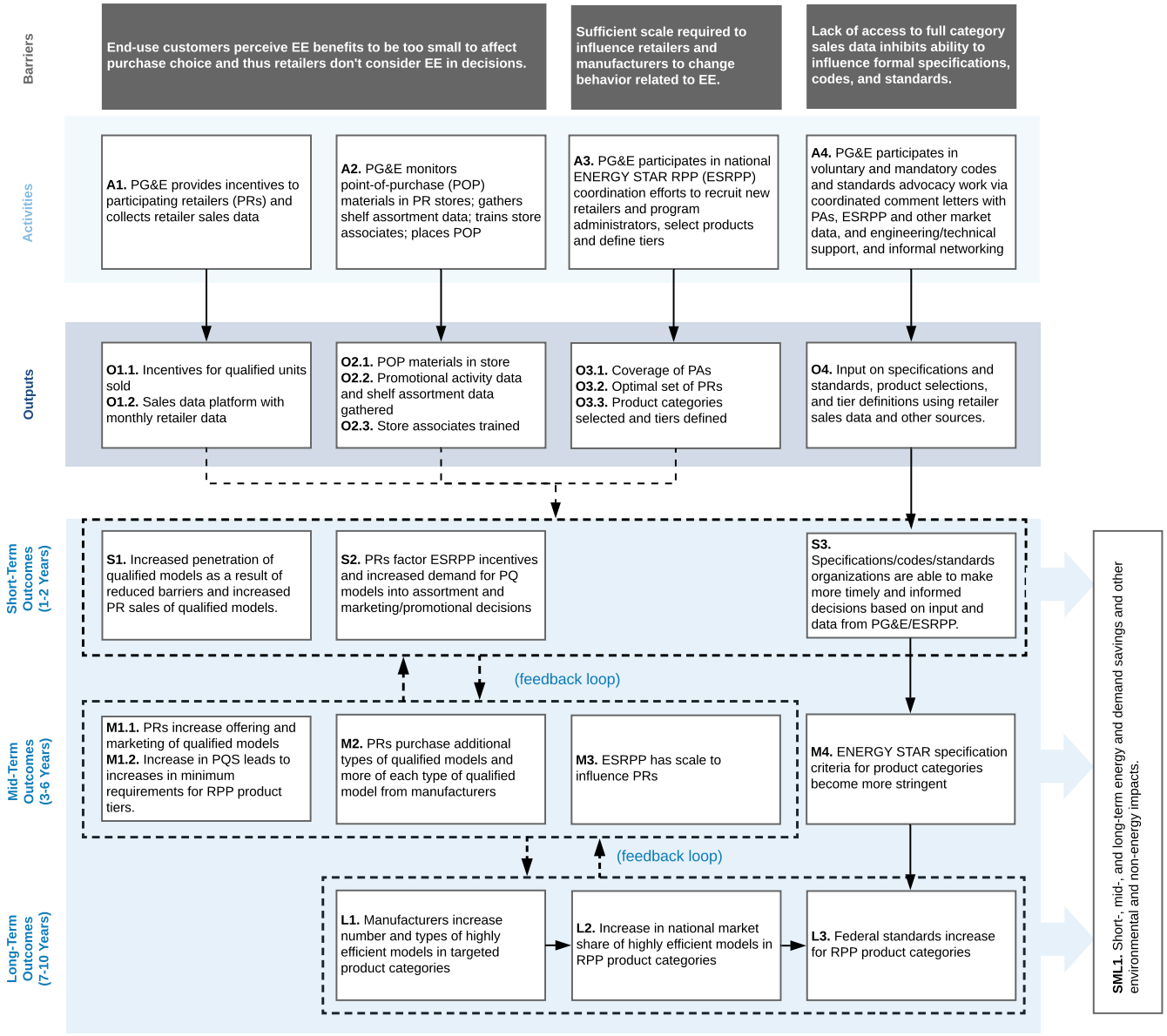
The evaluation team created a revised PG&E ESRPP logic model (presented below as Figure 3-1) to help guide subsequent research activities and frame the overall evaluation using the program theory. According to this program logic model, outcomes of the program are expected to evolve as the program increases in size and maturity:

- In the short-term (1-2 years), the program should gain sufficient scale to influence participating retailers' stocking and marketing, leading to increased sales of program-qualified models in participating store locations. At the same time, program delivery will be improved through the use of sales data and other information being tracked by program staff.

- In the mid-term (3-6 years), participating retailers should increase the proportion of qualified products in their assortment, begin to require more qualifying models from manufacturers, and favor program-qualified models in their marketing efforts. Program theory also suggests that energy efficiency criteria for qualifying products will increase, resulting in a "ratcheting up" of program eligibility requirements.
- In the long-term (7-10 years), manufacturers should increase the number and variety of energy-efficient models in targeted product categories, leading to a permanent increase in the availability of these models in retail stores, an increase in national market share for these models, and lastly, leading to more stringent federal standards for these product categories.

Validating the activities, outputs, outcomes, and linkages in this logic model allows us to assess the performance and efficacy of the program. Ultimately, this approach also allows us to inform estimates of program attribution. In Chapter 4, we provide a detailed assessment of program processes represented by the activities and outputs of the program logic model. In Chapter 5, we provide a detailed assessment of the short-, mid-, and long-term impacts.

Figure 3-1. Revised PG&E ESRPP Program Pilot Logic Model



Key: - - - - - - - **Dashed lines denote multiple interactions between elements inside (i.e., arrows connecting all elements).**

External Influences:
Broad economic conditions, market events, cost of energy, federal standards, ENERGY STAR, perceived need for conservation, and possible others.

Note: Factors can influence the program at all levels and time frames.

EVOLUTION OF PROGRAM THEORY

There have been several lessons learned since the PG&E ESRPP Program Pilot began operating in 2016. These lessons have altered the evaluation approach, and

informed the creation of the revised logic model shown above in Figure 3-1.[25] We discuss several of these lessons below.

- Overall program operations have been more fluid than initially envisioned, with a number of interactions occurring between elements in the program logic model. In the revised logic model (shown above as Figure 3-1) used for this evaluation research, these interactions are depicted by placing elements within dotted lines, rather than creating arrows from each element to the other. This depiction represents that within a dotted box, each element may potentially impact all other elements, resulting in a non-linear set of effects.
- During the early phases of program design, participating retailers were expected to commit to creating and implementing Retailer Implementation Plans ("Plans") for increasing the sales of energy-efficient models in the targeted product categories. These Plans would then serve as a tool to understand how retailers were using incentive dollars to drive sales of program-qualified units. In the course of this evaluation, it became clear that the Plans provided by retailers did not contain the level of detail initially expected by evaluators, and that there is no mechanism to obtain more-specific Plans. To address this, the evaluation approach has shifted slightly to place more weight on data collected from retailer store locations during in-store field visits by the PG&E ESRPP field services team.
- While the importance of the full category sales data collected from participating retailers has always been recognized, this pathway of influence within the program theory has become even more critical as our research has shown that such data simply does not exist elsewhere for the majority of the product categories included in ESRPP. As a result, the revised logic model more clearly emphasizes the importance of this data to the program's ability to facilitate the development of specifications, codes, and standards.

In the next section we provide a detailed explanation of the main levers that ESRPP uses to impact the market.

3.2 MAIN LEVERS OF ESRPP

In this section we provide clarification on the three main program levers of ESRPP: (1) the ability of ESRPP to work with participating retailers to connect

[25] Additional details showing how the revised logic model maps to the original logic model are provided in Appendix F.

manufacturers (further up the supply chain) with end-use customers (further down the supply chain), (2) the ability of ESRPP to collect and aggregate sales data from retailers, and (3) the ability of ESRPP to influence the development of codes, standards, and specifications. We discuss each of these points in more detail below.

RETAILERS ARE A CONDUIT CONNECTING MANUFACTURERS AND CUSTOMERS

While the PG&E ESRPP Program Pilot works with retailers to alter the market for energy-efficient models in select product categories, in the ESRPP program design, retailers are considered to be a conduit that connects manufacturers and end-use customers. By intervening in the middle of the supply chain, ESRPP seeks to overcome a "Catch-22" scenario whereby end-use customers do not consider the benefits of energy efficiency in their purchase decision, and thus retailers do not choose to stock or promote energy-efficient models. The effect of a $20 incentive paid to an end-use customer is likely insufficient to sway one's decision to purchase an energy-efficient model (as opposed to a standard efficiency model); however, when these incentives are paid to retailers *at scale*, program advocates believe that ESRPP can utilize retailers to alter customer decision-making through a number of possible mechanisms (including, but not limited to, activities that retailers do on a regular basis: advertising, assortment, product placement in stores, offering sale prices, and training employees to promote energy-efficient models).[26] A key long-term goal of ESRPP is to facilitate this movement towards greater efficiency, ultimately allowing the signal to be transmitted from end-use customers, through retailers, up to the manufacturers and suppliers responsible for producing the models.

FULL CATEGORY SALES DATA ARE A UNIQUE RESOURCE FOR BOTH EVALUATION AND ADVOCACY

In conducting research on sources of third-party market data with which to compare ESRPP sales data, the evaluation team determined that model-level sales data was unavailable for all but one product category (soundbars) currently included in the ESRPP portfolio. Some third-party data—such as the Association of Home and Appliance Manufactures (AHAM) shipment data used in this research—do provide total shipment data for some product categories at the monthly level. However, without model-specific values, it is impossible to compute program-

[26] It is important to note that in this program design, the end-use customer may be unaware that his or her purchase decision was altered through one of these mechanisms.

qualified share, which is one of the key metrics associated with gauging ESRPP progress toward market transformation.[27] This reality has added complications to any long-term evaluation of ESRPP, but it has also reinforced the importance of the full category sales data obtained by ESRPP sponsors from participating retailers. As discussed in more detail later in this report, this data is highly valued by entities like ENERGY STAR and has the potential to be very useful in advancing voluntary specifications for ESRPP product categories.

PG&E ESRPP CONTRIBUTIONS TO SPECIFICATIONS, CODES, AND STANDARDS ADVOCACY

Voluntary specifications (i.e., ENERGY STAR specifications), as well as mandatory efficiency standards (i.e., federal or state codes and standards), play a critical role in the ESRPP Program. Not only do ENERGY STAR specifications help to define the incented models within targeted product categories, mandatory codes and standards serve as baselines for estimating unit energy savings (for categories that are regulated by state or federal bodies). Additionally, because of the influence that the ESRPP program is expected to have on accelerating adoption of new voluntary specifications and mandatory codes and standards, some of the energy savings associated with these shifting specifications/standards will likely be attributable to the program. However, the evaluation of shifting codes and standards is a costly and complicated endeavor beyond the scope of this evaluation. If codes and standards for any of the ESRPP program products do change, it is expected that a parallel evaluation effort aimed at assessing the impacts of these changes on unit energy consumption and savings will be needed.

In the following section of this report, we provide an assessment of PG&E ESRPP Program Pilot processes.

[27] In discussions with AHAM in November 2018, it became apparent that AHAM is reevaluating the types of shipment reports it will publish, as one major manufacturer decided to no longer report shipment sales.

4. ASSESSMENT OF PG&E ESRPP PROCESSES

Through a program documentation review and interviews with both program staff and external collaborators, the evaluation team assessed key PG&E ESRPP program processes. In this chapter we provide an assessment of PG&E ESRPP Program Pilot processes. We first provide a high-level summary of this assessment, and then discuss results by individual activity and output. A detailed list of Program Performance Indicators is provided in Appendix D.

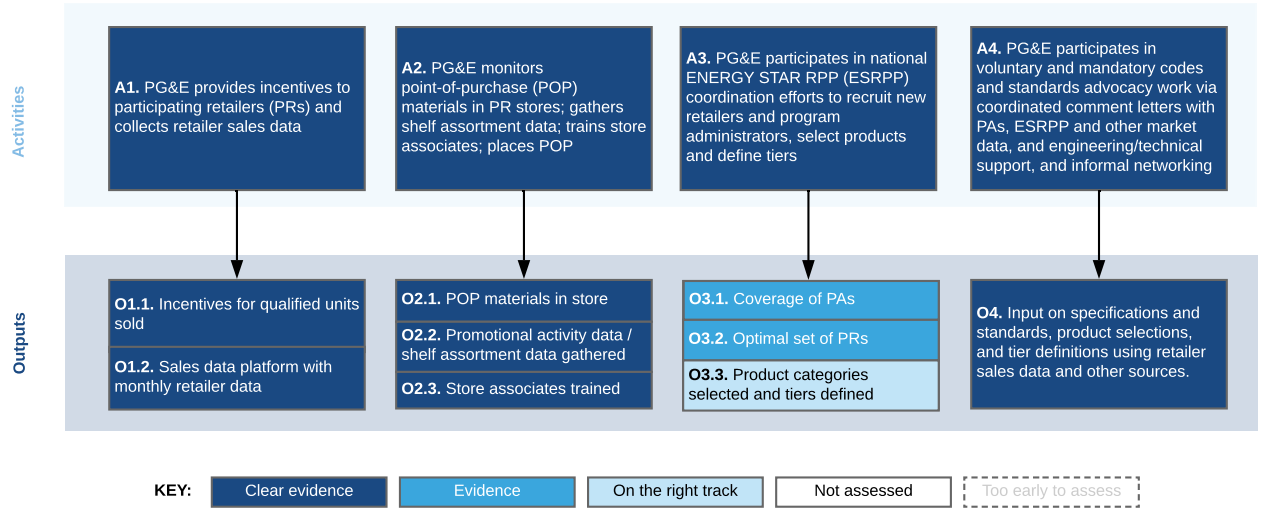
4.1 SUMMARY

The program pilot faced some major early hurdles related to data tracking and availability. Initially, challenges with the processing of retailer sales data made it difficult to perform thorough analysis of qualified models over time. This was more problematic for certain product categories where third-party data on the operating characteristics for individual models did not exist (air cleaners and soundbars). This processing has since become substantially more refined. Another early challenge was related to marketing plans that retailers were required to submit to the ESRPP program detailing their marketing plans for the upcoming program year. The idea is that these "retailer implementation plans" would enable evaluators to tie changes in sales back to specific activities undertaken by retailers and serve as evidence of attribution. However, in the course of early program development, it became clear that retailers were unable to provide this information in a format that would allow for the necessary analyses (likely because such marketing and promotional activities are not planned a year in advance). Instead, the program evaluation had to rely on in-store field data collected by a PG&E subcontractor to better understand what retailers were doing to promote and market program-qualified models.

At this point in time, PG&E ESRPP Program Pilot processes are generally working well, as reported by staff and external collaborator interviewees. As shown in the color-coded revised logic model shown in Figure 4-1, all activities and outputs are occurring as intended. The incentive payments to participating retailers and the subsequent collection of sales data—collectively the core "engine" of ESRPP—are occurring (logic model elements A1 and O1, as shown in Figure 4-1). Similarly, the collection of in-store field data and placement of ESRPP signage by PG&E has been successfully implemented (elements A2 and O2). PGE has successfully coordinated with the national ESRPP collaborative to recruit new retailers and select an optimal set of participating retailers, though some interviewees believe that the addition of more program administrators is necessary going forward (in order to achieve greater program scale). The definition of product eligibility tiers (element O3.3) is one area where program processes could be improved, as interviewees reported some challenges related to how ESRPP should define the appropriate levels. Lastly, there is clear evidence that PG&E's participation in activities related to specifications, codes, and standards (elements A4 and O4) is working well and leading to the intended outputs.

Staff interviewees indicated that they are satisfied with the ESRPP program because it has (1) gained recognition for its innovative approach to market transformation, (2) shown progress and results, and (3) received broad support from market actors. Some interviewees expressed dissatisfaction in the amount of progress related to measuring market transformation savings and progress toward achieving greater program scale, as well as lack of sufficient communication to all market actors so they can understand program developments.

Figure 4-1. Graphical Assessment of Activities and Outputs



In Table 4-1, we provide a high-level summary of the activities includes in the PG&E ESRPP logic model. We then provide a more detailed assessment of each activity.

Table 4-1. Summary of Evidence for Activities and Outputs

Logic Model Element	Summary
A1. PG&E provides incentives to participating retailers (PRs) and collects full-category retailer sales data	**SUPPORTING EVIDENCE:** Program data shows that PG&E is paying incentives to participating retailers on time and is successfully collecting and aggregating the resulting sales data. **SHORTCOMINGS:** Processing of retailer sales data is inherently difficult and must be carefully monitored.
A2. PG&E monitors point-of-purchase (POP) materials in PR stores; gathers shelf assortment data; trains store associates; places POP	**SUPPORTING EVIDENCE:** Field reports from PG&E's field services subcontractor detail efforts by PG&E to monitor POP materials placed by retailers, to gather shelf stocking/assortment data, train store employees, and place additional POP on program-qualified models. Staff interviews indicate that on the whole, these processes are operating smoothly. The evaluation team confirmed the field staff trained a total of 20,128 store associates from May 2016 - March 2018, with an increasing number of associates trained each quarter.
A3. PG&E participates in national ENERGY STAR RPP (ESRPP) coordination efforts to recruit new retailers and program administrators, select products and define tiers	**SUPPORTING EVIDENCE:** PG&E staff take frequent part in national coordination efforts through periodic conference calls, in-person meetings, and other forms of direct communication. PG&E staff also actively represent ESRPP's outreach efforts at conferences and industry trade shows. Along with the Northwest Energy Efficiency Alliance (NEEA), PG&E is considered by external collaborators to be highly influential within the ESRPP national effort. **SHORTCOMINGS:** The selection of products and the definition of qualifying tiers across all sponsors at the national level is a challenging aspect of the pilot program. Navitas (for PG&E) is working on creating a product portfolio and optimizing that portfolio. PG&E put together strategy documents for deciding when it makes sense to bring in a new product, but since product decisions are made at the national level, PG&E has limited ability to control this process.
A4. PG&E participates in voluntary and mandatory codes and standards advocacy work via coordinated comment letters with PAs, ESRPP and other market data, and engineering/technical support, and informal networking	**SUPPORTING EVIDENCE:** Overall, interview results show that while ESRPP is a national, collaborative program between many sponsors, PG&E is seen by external collaborators as a driving force behind the program. External collaborators also stated that overall ESRPP program success should be attributed to Program Administrators that were involved early on in the ESRPP program (of which PG&E was one). One external collaborator interviewee went as far as to say the fact EPA is supporting ESRPP is entirely due to PG&E, and now the program is a priority initiative at the EPA. Qualitative evidence from suppliers interviewed as part of this research suggests that the market actors are aware of and tracking ESRPP developments. **SHORTCOMINGS:** Careful documentation is required to show that voluntary standards/specification advocacy work is attributable to ESRPP and not other PG&E efforts.

4.2 ACTIVITIES AND OUTPUTS

In this section we provide additional details on findings that support our assessment of PG&E ESRPP Program Pilot activities and outputs. It is important to note that these analyses were designed to account for the fact that ESRPP operates on individual product categories. At the same time, we looked across product categories when providing an overall assessment of ESRPP program processes. These descriptions are organized by logic model element. A more detailed list of program performance indicators is provided in Appendix D.

PG&E PROVIDES INCENTIVES TO PARTICIPATING RETAILERS AND COLLECTS RETAILER SALES DATA (A1)

The process of paying incentives to retailers and collecting sales data generally works well, though there is some evidence that specific subprocesses could be fine-tuned. PG&E staff are currently working to develop program guidelines that can be shared with all ESRPP Program Administrators to help standardize some of the administrative processes that occur (e.g., payments to retailers). Interviewees included in this evaluation research noted that not all program administrators pay incentives at the same time, which could present a longer-term headache for participating retailers. One recommendation for standardization is for the administrators to not reconcile sales down to each store before paying retailers, since this practice can cause delays. PG&E currently reconciles retroactively and has facilitated prompt incentive payments.

Sales data continues to be a critical piece of ESRPP. Interviewees noted that it was due to one PG&E staff member's relationship with the participating retailers that enabled ESRPP to obtain highly sensitive full category sales data. External collaborators concurred that the full category sales data provided by retailers as part of the ESRPP program requirements was a notable achievement. Interviewees noted their satisfaction with participating retailers' support, especially given they now provide that full category sales data.

PG&E MONITORS POINT-OF-PURCHASE (POP) MATERIALS IN PARTICIPATING RETAILER STORES; GATHERS SHELF ASSORTMENT DATA; TRAINS STORE EMPLOYEES; PLACES POP (A2)

Based on PG&E staff interviews, PG&E program staff work closely with subcontractors to perform several key field activities in participating retail stores, including the collection of shelf survey data and the placement of point-of-purchase (POP) materials. The subcontractor field staff visit participating retail stores in PG&E service territory each month to place POP materials and conduct "baseline" shelf surveys on a subset of store visits. Field staff reported that while these surveys

were initially conducted using pen and paper, they had transitioned to a digital version which helped speed up the process.

PG&E subcontractors mentioned they are communicating essential program knowledge in the field, such as the field subcontractor providing training on program qualified products to sales representatives and another subcontractor communicating feedback and installation guidance on marketing collateral.

PG&E PARTICIPATES IN NATIONAL ESRPP COORDINATION EFFORTS TO RECRUIT NEW RETAILERS AND PROGRAM ADMINISTRATORS, SELECT PRODUCTS, AND DEFINE PRODUCT TIERS (A3)

Evaluation research shows that PG&E staff actively represent ESRPP's outreach efforts through multiple forms. Staff report presenting at conferences, sitting on professional panels, and attending trade shows to discuss and promote the ESRPP program in an effort to recruit new program administrators.[28] Nationally, ESRPP currently has 14 utilities in 15 states, representing roughly 18% coverage of the US population; however, the program is still short of its goal of 30% national coverage.

According to external collaborators interviewed as part of this evaluation, ESRPP has had a smooth retailer recruitment process overall. Currently, the program is not looking to add any new retailers until an evaluation occurs. However, PG&E has identified additional retailers who are potentially interested in participating in the near future.

External collaborators interviewed as part of this evaluation also reported that PG&E is considered to be highly influential within the ESRPP national effort, along with the Northwest Energy Efficiency Alliance (NEEA).

One difficult aspect of implementing the ESRPP program is the selection of products and the definition of qualifying tiers across all sponsors at the national level. While documentation for these processes does exist, each product category is unique in some way, making it difficult to standardize the process across the entire portfolio of products. There are several important aspects to this:

- PG&E and other program administrators must decide when it makes sense to either bring a new product into the program or remove an existing product from the program. For instance, ESRPP makes significant decisions on target product categories, such as postponing including TVs in the program. PG&E and other sponsors look at market penetration of different

[28] A more detailed list of PG&E ESRPP outreach and advocacy efforts is included as Appendix E.

levels of ENERGY STAR and ENERGY STAR Most Efficient, as well as how many products are available to incent.

- Program sponsors also determine what level of incentive, per product, will incentivize the retailers to ask manufacturers to provide more of them.

There is some disagreement between interviewees regarding the national ESRPP approach to setting program qualifications requirements, with some interviewees indicating that the program requirements should be simplified while other interviewees supported the current program approach to setting tiers that are higher than current ENERGY STAR specification levels (e.g., ENERGY STAR + 50%). ***(It is important to note that product selection decisions are made at the national level, and thus PG&E has limited ability to control this process.)*** Multiple interviewees mentioned that ESRPP may want to reconsider the way it currently sets tiers above the existing ENERGY STAR specification level for products with relatively high market share of qualified models.

The initial impetus for this adjustment was the need to "ratchet up" requirements for product categories where the base ENERGY STAR level had already achieved a substantial share of the market. But while interviewees felt this was the correct approach in theory, it was logistically problematic, as retailers have a difficult time understanding what is qualified and what is not qualified. One recommendation was to use an obvious feature or functionality to determine eligibility instead of an "ENERGY STAR + X %" requirement. Staff mentioned the possibility that paring down the number of product categories and/or tiers may help streamline program processes moving forward.

PG&E PARTICIPATES IN VOLUNTARY AND MANDATORY CODES AND STANDARDS ADVOCACY WORK VIA COORDINATED EFFORTS (A4)

Overall, interview results show that while ESRPP is a national, collaborative program between many sponsors, PG&E is seen by external collaborators as a driving force behind the program, particularly for efforts aimed at advancing voluntary specifications. Specific supporting evidence includes the following:

- One external collaborator interviewee went as far as to say the fact EPA is supporting ESRPP is entirely due to PG&E and now the program is a priority initiative at the EPA.
- Multiple external collaborators reported that the full category sales data has been an important tool in specification-setting efforts and has allowed them to drive specifications higher than they otherwise would have been able to do.
- One external collaborator interviewee noted that ESRPP has "provided good conversations" around qualifying product levels and national analysis of market share. In particular, this interviewee stated the work so far has "pushed the envelope on efficiency standards" by supplying supporting data

collected from retailers and noted, as an example, that the EPA put out a discussion guide on air cleaners that was prompted by ESRPP's market data.

- One interviewee was quoted saying, "Specification-setting runs the risk of not having a balanced pool if utilities are not accounted for. [ESRPP] helps EPA defend against pushback from manufacturers and others."

PG&E ESRPP program staff work with PG&E's internal Codes and Standards group to communicate with mandatory codes and standards-setting organizations such as the California Energy Commission (CEC) and DOE. Interviewees reported that it is important for the different teams at PG&E to coordinate internally to make sure the messaging is consistent across these efforts and that the same data is being used where appropriate. However, the boundaries between ESRPP and other codes and standards efforts become fuzzy in some of these instances, making it more difficult to clearly delineate the role of ESRPP.

A more detailed documentation of PG&E ESRPP outreach and advocacy efforts is included in Appendix E.

5. ASSESSMENT OF PG&E ESRPP IMPACTS

Impacts for a market transformation program like ESRPP take different forms. The primary impacts examined as part of this early evaluation effort included:

- Increases in sales and program-qualified market share among participating retailers (logic model element S1)
- Increases in the proportion of program-qualified models on store shelves (logic model element M1.1)
- Improvements in the ability of organizations working on specifications, codes, and standards to make better decisions based on inputs from ESRPP (logic model element S3)

These analyses were designed to account for the fact that ESRPP operates on individual product categories. At the same time, we looked across product categories when providing an overall assessment of ESRPP program impacts.

In this chapter we provide an assessment of PG&E ESRPP Program Pilot impacts to date. We first provide a high-level summary of this assessment, and then discuss results by time period (short-, mid-, and long-term). A more detailed list of market transformation indicators is provided in Appendix D.

5.1 SUMMARY

Early evaluation results provide evidence that the PG&E ESRPP program is leading to short-term and mid-term impacts as expected by program theory, though the results differ by product category, and results are not uniformly positive. Based on statistical modeling of retailer sales data, we observe short-term sales increases for 5 of 11 tiers (covering 4 of 7 product categories) currently targeted by ESRPP (logic model element S1). At the same time, interviews with national-level retail staff show that ESRPP incentives have some influence on retailer decision-making, and interviews with external collaborators show that ESRPP is facilitating the development ENERGY STAR specifications. At this point in time, it is premature to assess the long-term outcomes included in the logic model.

There are also a number of shortcomings identified in the program pilot's ability to achieve its desired short-term and mid-term outcomes.

- In looking at short-term outcomes, we do not see statistically-significant increases in sales across *all* product categories and tiers receiving incentives (logic model element S1).[29]
- Secondly, although retailers do indicate that ESRPP incentives have factored into their decision-making (logic model element S2), it remains difficult to understand exactly how the ESRPP incentives are considered by retailers relative to other considerations (e.g., manufacturers competing for shelf space).
- Lastly, at this stage, much of PG&E ESRPP's advocacy efforts have been aimed at ENERGY STAR, with limited activities aimed at other standards-setting bodies (logic model element S3).

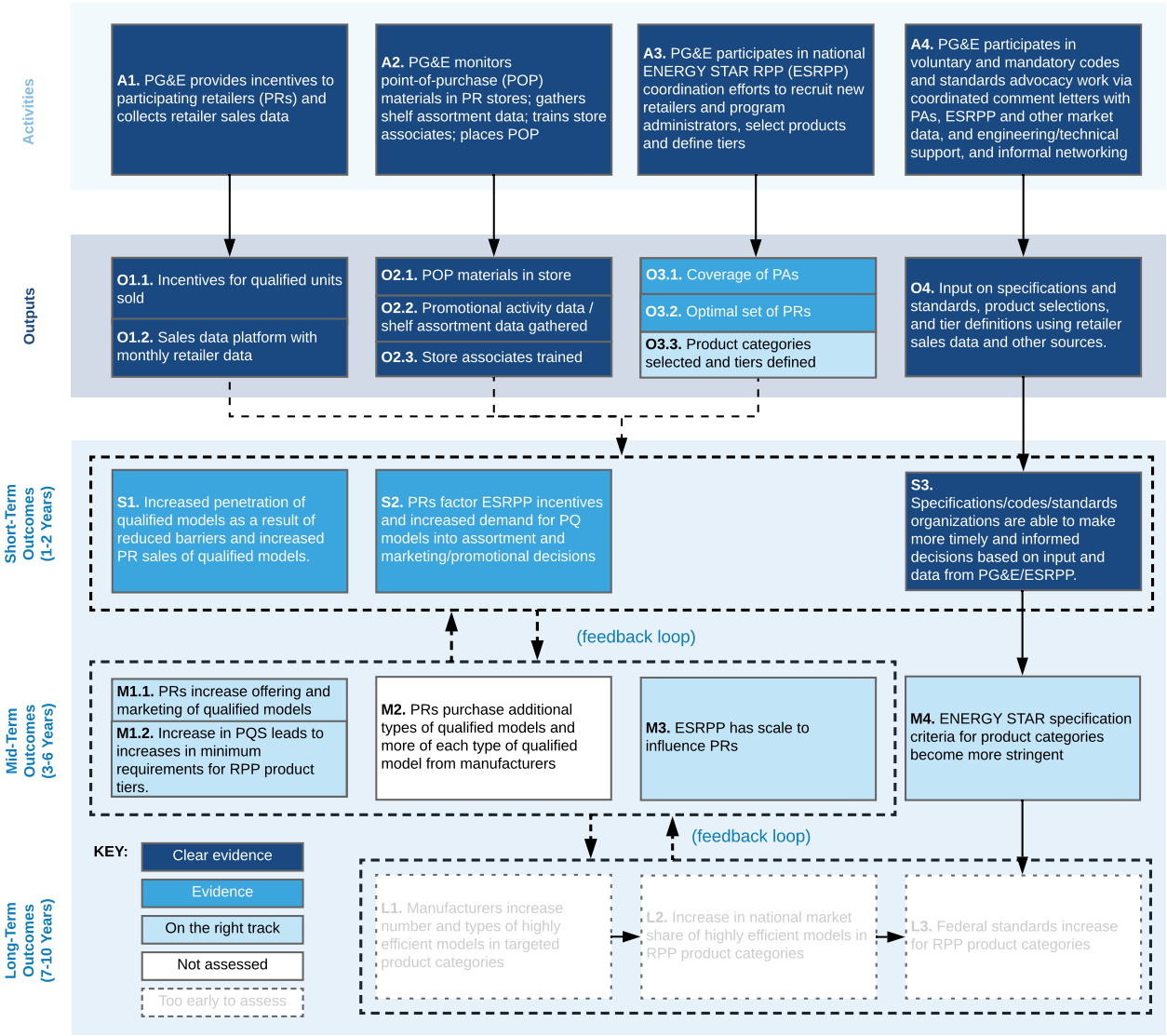
In looking at mid-term outcomes, we see several additional shortcomings.

- While our analysis of shelf survey data shows increases for 5 of 7 product categories, an additional two product categories are either flat (air cleaners and soundbars) or have a decreasing trend in model assortment share. The reasons for this are unclear. However, because PG&E has identified that the primary objective for these categories is to facilitate the advancement of ENERGY STAR specifications through the provision of market data, the lack of increasing market share for these categories becomes less important.
- In the mid-term, increasing market share is designed to lead to a "ratcheting up" of program requirements. In many cases this requires setting an efficiency level based on ENERGY STAR, but does not map directly to an existing designation such as ENERGY STAR Most Efficient. This new level instead takes the form of "ENERGY STAR + XX%" and makes it difficult for retailers (and potentially customers) to easily understand which models are program-qualified.

We provide a graphical depiction of progress toward expected outcomes below in Figure 5-1.

[29] As noted elsewhere in this report, for some product categories/tiers, increasing sales of program-qualified models was not necessary an objective.

Figure 5-1. Graphical Summary of PG&E ESRPP Impacts



In the following sections we provide additional details on findings that support our assessment of PG&E ESRPP Program Pilot impacts. We first discuss short-term outcomes, then mid-term outcomes, and finally long-term outcomes.

5.2 SHORT-TERM OUTCOMES

In the short-term (1-2 years), the PG&E ESRPP Program Pilot is expected to result in three key outcomes:

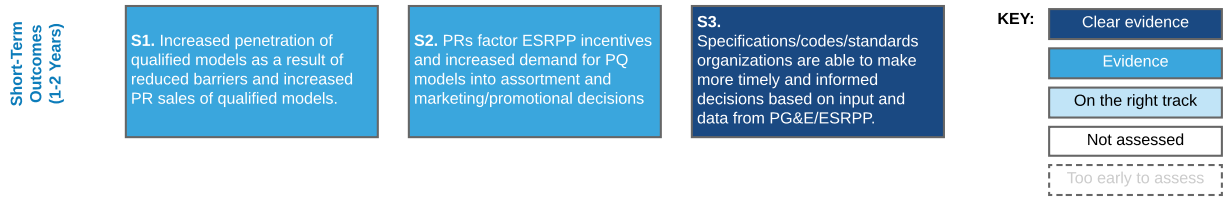
- Sales of program-qualified models should increase for participating retailers, leading to a corresponding increase in program-qualified market share (logic model element S1).

- Participating retailers should begin to consider this sales increase in their marketing and promotional decisions (logic model element S2).
- Organizations responsible for creating specifications, codes, and standards should be able to make more timely and informed decisions (largely due to the availability of the full category sales data obtained through ESRPP) (logic model element S3).

There is evidence from the evaluation research that the PG&E ESRPP Program Pilot has made significant progress toward all three of these short-term outcomes in the program logic model. As shown in Figure 5-2, there is evidence of increased sales and program-qualified market share for some product categories (logic model element S1). There is also evidence from retailer interviews that the ESRPP incentives have been factored into retailer decision-making (logic model element S2). Lastly, there is clear evidence from interviews with external collaborators that ESRPP has empowered programs like ENERGY STAR to improve their decision-making regarding specification revisions for product categories includes in ESRPP.

There are also a number of challenges associated with the short-term outcomes. For one, we do not see statistically-significant increases in sales across all product categories and tiers receiving incentives.[30] Secondly, although retailers do indicate that ESRPP incentives have factored into their decision-making, it remains difficult to understand exactly how the ESRPP incentives are considered by retailers relative to other considerations (e.g., manufacturers competing for shelf space). Lastly, at this stage, much of PG&E ESRPP's advocacy efforts have been aimed at ENERGY STAR, with limited activities aimed at other standards-setting bodies.

Figure 5-2. Short-Term Outcomes



In the sections below we highlight specific evidence for this assessment of short-term outcomes.

[30] As noted elsewhere in this report, for some product categories/tiers, increasing sales of program-qualified models was not necessary an objective.

Table 5-1. Summary of Assessment of Short-Term Outcomes

Logic Model Element	Summary
S1. Increased penetration of qualified models as results of reduced barriers and increased PR sales of qualified models	**SUPPORTING EVIDENCE:** Results from the sales data modeling show a statistically significant increase in sales for 5 of the 11 tiers (or 4 of the 7 product categories) incented by PG&E ESRPP. **SHORTCOMINGS:** Increases in sales share are not detectable in all product categories/tiers incented. In some cases it is not readily apparent why the intervention works for some categories but not others. For some categories, limitations in modeling present difficulties when estimating the baseline, particularly for seasonal categories like room air conditioners.
S2. Participating retailers factor incentives and increased demand for program-qualified models into assortment and marketing/promotional decisions	**SUPPORTING EVIDENCE:** Senior management at participating retailers believe the ESRPP program influences their pricing decisions to some degree; however, for ESRPP to be a primary driver in decision-making, they believe the program needs to "scale up." **SHORTCOMINGS:** Despite this feedback from retailers, it remains difficult to directly gauge ESRPP's level of influence on retailer decision-making. In a highly-competitive industry with slim margins and quickly-changing customer preferences, retailers are hesitant to provide any information related to promotional or marketing strategies. This is particularly true for the requested retailer implementation plans, which did not contain sufficient detail to track marketing activities. Thus the program has had to rely more on in-store data collection activities rather than rely on retailer marketing plans.
S3. Specifications/Codes/Standards Organizations are able to make more timely and informed decisions based on input and data from ESRPP	**SUPPORTING EVIDENCE:** External collaborators indicated that the full category sales data provided by retailers as part of the ESRPP program requirements was a notable achievement. EPA staff reported that these data have been an important piece in specification-setting efforts and have allowed them to drive specifications higher than they otherwise would have been able to do. External collaborators noted that PG&E has been able to leverage its own internal resources to address technical testing needs related to two ESRPP products (soundbars and air cleaners) and that the results of this testing have also aided specification-setting efforts. **SHORTCOMINGS:** The majority of PG&E ESRPP's direct influence appears to be on EPA/ENERGY STAR. Additional influence on other codes/standards bodies (such as DOE) is significantly less, at least at this stage. The relationship with federal bodies is additionally complicated by political dynamics outside the control of PG&E ESRPP.

In the following sections we provide more details on each of these outcomes.

INCREASED SALES AND PENETRATION OF PROGRAM-QUALIFIED MODELS (S1)

Results from the sales data modeling vary by product category and tier:

- We observe a statistically significant increase in sales for 5 of 11 tiers currently incented by the PG&E ESRPP program pilot.
- For an additional 3 of 11 tiers (basic air conditioners, advanced air conditioners, and advanced freezers) we were unable to estimate any statistically significant changes due to the ESRPP program, either due to a small number of data points or extreme seasonality in the sales stream.
- For an additional 3 of 11 tiers currently receiving incentives (basic air cleaners, advanced air cleaners, and advanced washers), we do not see any corresponding increase in sales.[31]

These results are summarized below in Table 5-2.

[31] One of these tiers (basic soundbars) was incented only in PY 2016. No increase in sales was observed for this tier.

Table 5-2. Summary of Sales Increases by Product Category and Tier

Product Category	Tier	Years Incented		Increase in Sales Above Baseline?	Notes
		PY 2016	PY 2017		
Air Cleaners	Basic	Yes	Yes	No	
	Advanced	Yes	Yes	No	
Air Conditioners	Basic	Yes	Yes	Indeterminate	Substantial uncertainty in modeling due to extreme seasonal sales fluctuations.
	Advanced	No	Yes	Too few sales	
Dryers	Basic	Yes	Yes	Yes	
	Advanced	Yes	Yes	Yes	
Freezers	Basic	Yes	Yes	Yes	
	Advanced	Yes	Yes	Too few sales	
Refrigerators	Basic	No	No	No	
	Advanced	No	Yes	Yes	
Soundbars	Basic	Yes	No	No	There is the possibility of "cannibalization" from the basic tier to the advanced tier.
	Advanced	Yes	Yes	Yes	
Washers	Basic	No	No	Yes	The difference between qualifying levels for basic and advanced is very small (5%). The overall product category shows a statistically significant increase.
	Advanced	No	Yes	No	

We summarize each of these trends below by product category, and where possible, provide additional context to help explain these findings:

- For air cleaners, we found decreases for both the basic and advanced tiers, but do not believe these decreases are due to the program. In particular there was a massive decrease in program-qualified sales share around the time that wildfires made air quality in the PG&E service territory very bad, and while unit sales of qualified units increased, the decrease in program-qualified share led to an overall decrease in predicted sales.
- For air conditioners, the differences were not statistically significant, likely due to the fact that the seasonality leads to such little pre-program data. The analysis of room air conditioners was also complicated by atypically hot temperatures in much of California in 2017 (compared to 2016).
- For dryers we observed an increase in program-qualified sales for both the basic and advanced tiers.
- For freezers we observed an increase in the basic tier as well as for all qualified products. We did not see an increase for advanced freezers, though because the sales were so small for advanced freezers in the pre-program period and essentially zero in the program period, this result is of limited practical importance.
- For refrigerators, we observed an increase in sales in the advanced tier. (PG&E does not incent basic tier models, and the change for basic tier models was not statistically significant.)

- Soundbars exhibited a shift from the basic tier to the advanced tier (i.e., there was a decrease in sales in the basic tier and an increase in sales in the advanced tier); we consider both changes to be due to the program to avoid over-counting impacts from the increases in the advanced tier, some of which are likely due to cannibalization between tiers in addition to overall increases in the advanced tier.
- For washers, we observed a statistically significant increase for basic tier models but not advanced tier models. Given that incentives were provided for the sales of advanced tier models but not basic tier models, this result is counterintuitive. However, we note that the difference in program-qualifying requirements between the basic and advanced tiers for washers is very small (ENERGY STAR vs. ENERGY STAR + 5%). It is possible that retailer efforts to increase the sales of advanced tier models (i.e., ENERGY STAR + 5%) may have also had an effect on basic tier models, even though these models did not themselves receive incentives.

Specific results by product category and tier are presented below in Table 5-3.

Table 5-3. Increased Sales (Above Baseline) of Program-Qualified Models

Product	Tier	Increase	SE	Lower Bound (95% CI)	Upper Bound (95% CI)	PQ Sales in the Post Period	% Change
Air Cleaners	Basic	-1,207	300	-1,716	-698	22,472	-5%
Air Cleaners	Advanced	-3,088	201	-3,429	-2,748	4,052	-76%
Air Cleaners	Qualified	-2,949	293	-3,446	-2,452	26,737	-11%
Room Air Conditioners	Basic	-6,983	6,973	-18,807	4,841	59,288	-12%
Room Air Conditioners	Qualified	-1,902	5,242	-10,790	6,987	59,813	-3%
Dryers	Basic	10,557	2,483	6,347	14,766	135,388	8%
Dryers	Advanced	159	20	125	193	279	57%
Dryers	Qualified	9,729	2,477	5,529	13,928	135,692	7%
Freezers	Basic	2,523	352	1,925	3,120	17,057	15%
Freezers [a]	Advanced	-5,357	761	-6,646	-4,067	1	-535669%
Freezers	Qualified	10,727	518	9,849	11,605	17,054	63%
Refrigerators	Basic	-3,142	3,215	-8,718	2,434	160,061	-2%
Refrigerators	Advanced	6,933	1,056	5,102	8,764	45,948	15%
Refrigerators	Qualified	3,075	2,780	-1,745	7,895	209,126	1%
Soundbars [b]	Basic	-986	61	-1,089	-883	421	-234%
Soundbars	Advanced	10,477	116	10,281	10,673	15,057	70%
Soundbars	Qualified	10,045	114	9,851	10,238	15,603	64%
Washers	Basic	92,754	1,624	89,938	95,569	108,917	85%
Washers	Advanced	-93,387	6,006	-103,802	-82,972	89,450	-104%
Washers	Qualified	7,106	2,138	3,399	10,813	202,395	4%

[a] A small number of sales in the program period for advanced freezers leads to large relative changes.
[b] PG&E stopped incenting basic tier soundbars in 2017.

Although PG&E does not currently plan to claim short-term savings associated with the program pilot, EMI Consulting did estimate energy and demand savings to understand what the magnitude of these savings might be. These results are shown in Table 5-4. As discussed in Section 2.4, energy savings are derived from average energy savings in the post-period for each tier, multiplied by the estimated sales increases for that product category and tier. We did not estimate savings for product categories/tiers without an increase in sales above baseline.

Table 5-4 Energy Savings for Program-Qualified Models for Program Years 1 and 2 (2016-2018)

Product [a]	Tier	Energy Savings (kWh)	Demand Reduction (kW)	Energy Savings (Therms)
Dryers	Basic	732,333 ± 292,022	123.561 ± 49.271	8,430 ± 3,362
	Advanced	23,262 ± 4,963	4.496 ± 0.959	-277 ± 59
Freezers	Basic	43,6241 ± 35,692	92.310 ± 7.552	-11,952 ± 978
Refrigerators	Advanced	480,971 ± 126,996	105.191 ± 27.775	-13,010 ± 3,435
Soundbars	Basic	-35,536 ± 3,718	-0.660 ± 0.069	822 ± 86
	Advanced	258,262 ± 4,831	4.020 ± 0.075	-5,902 ± 110

[a] Savings estimates with "+" values constitute a 90% confidence interval. Negative values represent negative savings due to interactive effects.
[b] PG&E stopped incenting basic tier soundbars in 2017.

INCREASED PROGRAM-QUALIFIED SALES FACTORED INTO RETAILER DECISION-MAKING (S2)

According to Cadmus' national-level year two interviews, internal stakeholders at participating RPP retailers believe that the key driver of product assortment process—which is also believed to be the key driver of increasing ESRPP product sales—is profitability. Since profitability is the most important factor to merchants, retailers disclosed they are largely distributing incentives to program-qualified product SKUs to make them appear more favorable and influence merchants' stocking decisions.

- These interviews showed that merchants and manufacturers have some influence on marketer's decisions regarding which products to promote, however, much of the retailer's marketing strategy is determined either at the national level or at the beginning of the year. Therefore, marketing staff make specific plans for ESRPP products but do not focus on it because their marketing focus is generally at a national level.

- According to Cadmus' national-level interviews with Nationwide in their first year of participation (PY 2017), senior management see ENERGY STAR as a point of differentiation for them, compared to other retailers, due to their level of commitment and the marketing support they put behind ENERGY STAR products. Nationwide was already a "strong performer" in selling ESRPP program-qualified models before joining ESRPP and claims to have seen a boost to sales since joining. Like the other participating retailers, Nationwide's primary decision factor is profitability. However, senior staff believe the biggest barrier to making profit from energy-efficient products is *lack of information*. These staff think ESRPP will help in training sales associates to sell program-qualified products to customers.

Despite this feedback from retailers, it remains difficult to directly gauge ESRPP's level of influence on retailer decision-making. In a highly-competitive industry with slim margins and quickly-changing customer preferences, retailers are hesitant to provide any information related to promotional or marketing strategies. Thus it becomes more critical to observe actual changes in retailer practices (for instance, by in-store data collection) rather than rely on prospective plans.

We note that there does appear to be a positive (albeit very qualitative) change in senior management views of ESRPP influence over time. One finding from the analysis of the 2016 interviews: "Senior management is starting to get interested, but because ESRPP is still small, the interest is limited." A finding from the 2018 interviews: "ESRPP is overall well received by retailers, including merchants, marketers, and VPs. Retail support for ESRPP has increased slightly since the first year."

SPECIFICATIONS/CODES/STANDARDS ORGANIZATIONS ABLE TO MAKE MORE TIMELY AND INFORMED DECISIONS (S3)

External collaborators believe that the process by which PG&E works with the ENERGY STAR program to advance ENERGY STAR-qualifying requirements works well. Interviewees noted that the specification development process is typically spread out over a long time period, and so the types of interactions between PG&E ESRPP staff and ENERGY STAR will depend on the schedule for draft specification revisions or new specifications.

PG&E ESRPP involvement with other codes and standards organizations, such as the CEC and DOE, take place primarily through PG&E's internal Codes and Standards group. Staff indicated that a key piece in codes and standards advocacy work is a "Code Change Theory Report" (CCTR), which details the "story behind the rule-making process." Interviewees noted that mandatory codes and standards (either state or federal) change much less frequently than do ENERGY STAR specifications, and so any ESRPP program impacts on such changes should be viewed through this lens. Interviewees also noted that the pace of progress for

federal standards in particular is dependent on national political dynamics and that, recently, there has been much less activity occurring than in previous years.

5.3 MID-TERM OUTCOMES

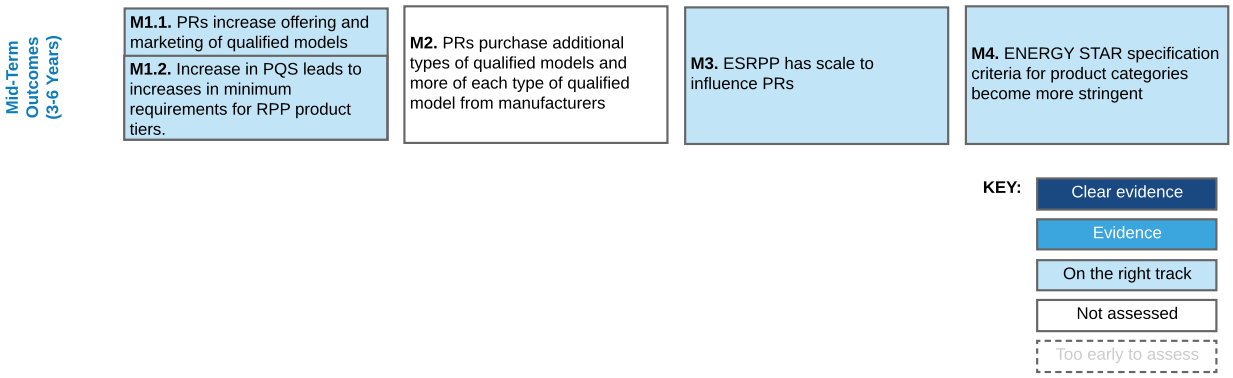
In the mid-term (3-6 years), the PG&E ESRPP Program Pilot is expected to result in five key outcomes:

- Participating retailers should increase their offering and marketing of program-qualified models, leading to an increase in the minimum qualifying requirements for ESRPP (logic model elements M1.1 and M1.2).
- To address the increase in demand, participating retailers should begin to purchase additional types of qualified models and more of each type from manufacturers (logic model element M2).
- As additional program administrators join the national collaborative, ESRPP should have achieved sufficient scale to increase retailer decision-making at the national level (logic model element M3).
- ENERGY STAR specification criteria for ESRPP product categories should become more stringent (largely due to the availability of the full category sales data obtained through ESRPP) (logic model element M4).

Evaluation results show the PG&E ESRPP Program Pilot is on track to achieve four of these mid-term outcomes in the program logic model (depicted graphically in Figure 5-3). We were unable to assess whether or not participating retailers are requesting additional qualified models/units from manufacturers (outcome M2) since the data required for this assessment have not yet been collected.[32]

[32] The data required for this will be based on interviews that Cadmus is slated to conduct with manufacturers in 2019.

Figure 5-3. Mid-Term Outcomes



In Table 5-5 and Table 5-6 below, we provide a summary of the specific evidence supporting this assessment.

Table 5-5. Summary of Assessment of Mid-Term Outcomes

Logic Model Element	Summary
M1.1. Participating Retailers increase offering and marketing of qualified models	**SUPPORTING EVIDENCE:** Analysis of in-store shelf assortment data collected during field visits shows an increasing trend in model assortment share for five product categories (dryers, refrigerators, room ACs, freezers, and washers). Analysis of this data also shows that retailers give preferential treatment to qualified models for all 9 of the 11 product tiers incented, though it is difficult to know what this would have looked like in the absence of ESRPP. **SHORTCOMINGS:** An additional two product categories are either flat (air cleaners and soundbars) or have a decreasing trend in model assortment share. The reasons for this are unclear.
M1.2 Increase in program-qualified share leads to increases in minimum requirements for ESRPP product tiers	**SUPPORTING EVIDENCE:** ESRPP eligibility requirements have "ratcheted up" for several products—including air cleaners and washers—in order to readjust for high market share. National requirements have also ratcheted up for the basic tier of refrigerators, though PG&E does not support currently this tier. Additionally, the program has made several additional adjustments to program requirements: (1) PG&E stopped incenting basic soundbars after PY 2016 due to high market share, and (2) PG&E decreased the incentive for basic tier dryers in PY2017. **SHORTCOMINGS:** "Ratcheting up" of program requirements in many cases requires setting an efficiency level based on ENERGY STAR, but which does not map directly to ENERGY STAR Most Efficient. This new level takes the form of "ENERGY START + XX%" and makes it difficult for retailers (and potentially customers) to easily understand which models are program-qualified.

Table 5-6. Summary of Assessment of Mid-Term Outcomes (continued)

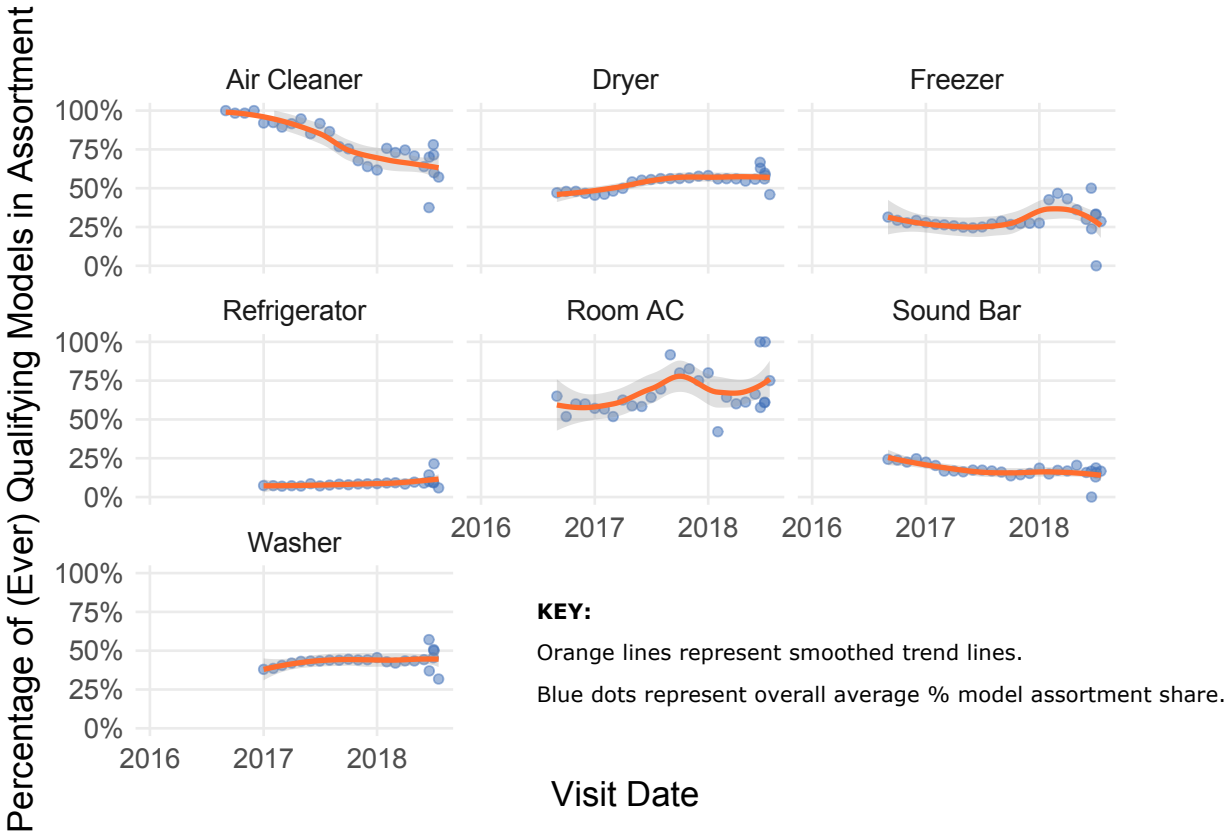
Logic Model Element	Summary
M3. ESRPP has scale to influence participating retailers	**SUPPORTING EVIDENCE:** Nationally, ESRPP has achieved coverage of an estimated 18% of the US population. This is lower than ESRPP's stated goal of 30% (based on retailer input), but high enough to keep retailers engaged with the program. Two suppliers interviewed as part of this research both provided confirmation that the ESRPP program has achieved sufficient scale such that market actors operating higher up in the supply chain (e.g., manufacturers and suppliers) are considering its impact on their business strategies, though the magnitude of the impact at this point is difficult to assess, and likely small. **SHORTCOMINGS:** Retailers indicate that the ESRPP program should attain a 30% market coverage nationally in order to be more effective. The national ESRPP program will require more Program Administrators to join before this is possible.
M4. ENERGY STAR specification criteria for product categories become more stringent	**SUPPORTING EVIDENCE:** Interviews conducted as part of this evaluation research suggest that ESRPP inputs have had a direct effect on the development of ENERGY STAR specification criteria, though not enough time has passed for these effects to be fully realized in the adoption of new specifications. Directional evidence for this assertion comes from interviews with external collaborators, who attest that the sales data and technical input from ESRPP has been a critical piece of the specification development process. Interview results suggest that PG&E ESRPP efforts have been particularly fruitful for air cleaners and soundbars.

In the following sections we provide more details on each of these outcomes.

PARTICIPATING RETAILERS INCREASE OFFERING AND MARKETING OF QUALIFIED MODELS (M1.1)

Program theory holds that as participating retailers are able to use incentive dollars to drive demand for energy efficient models, they must (in the mid-term) adjust their stock accordingly to address this. By looking at model assortment share on participating retailers' shelves over the course of the program period, we were able to assess whether this trend was observable. As shown below in Figure 5-4, we observed slight upward trends for five product categories (dryers, freezers, refrigerators, room ACs, and washers). We also observed two downward trends—for air cleaners and soundbars—though the data suggest these trends were driven largely by one or two retailers only.

Figure 5-4. Model Assortment Share Over Time[a]



a Trends shown here are for program period only, with the exception of refrigerators and washers, which began receiving incentives in April 2017.

In Table 5-7 we place these results in context with the results of the sales data modeling looking at increase in program-qualified sales share (a short-term outcome). This comparison shows that for most product categories, the results of the two analyses support a coherent storyline—that an increase in sales share of program-qualified models in the short-term may plausibly lead to an increase in program-qualified shelf assortment share in the mid-term. However, it is important to note that the two phenomena represent different measurements of program impacts, and do not necessarily align with each other.

Table 5-7. Comparison of Sales Data Analysis Results and Shelf Assortment Analysis Results

Product Category	Program-Qualified Sales Increase?	Program-Qualified Shelf Assortment Increase?
Air Cleaners	No	No
Dryers	Yes**	Slight increase**
Freezers	Yes (basic tier only)**	Slight increase**
Refrigerators	Yes (advanced tier only)**	Slight increase**
Room ACs	Indeterminate	Slight increase**
Soundbars	Yes (advanced tier only)**	No
Washers	No	Slight increase**

All increases were statistically-significant, p<.05.

In addition to stocking and assortment, program theory indicates that availability of incentives will lead retailers to provide qualified products with preferential treatment in their internal promotion decisions. As discussed in Section 3.1, the original evaluation plan envisioned using Retailer Implementation Plans to assess how the program was influencing these internal decisions, but in reality, the retailer implementation plans lack sufficient specificity to assess promotional decisions. In lieu of this data source, we have reviewed data collected by PG&E's field services subcontractors regarding the placement and pricing of products at participating retailers' stores.

We relied on model-level data collected from 403 individual store visits across 288 retail locations between January and August 2018. We calculated the percentage of models that were placed in a preferential location (anything other than simply in the aisle), the percentage of models that were currently discounted relative to the regular price, and the average discount amount among discounted products by product group and tier (non-qualified, basic, and advanced, as well as all qualified). The qualified designation is based on the categorization conducted by the field team; the basic and advanced tier designations are based on model matching we conducted for this analysis.

Results of this analysis are shown in Table 5-8. We found that all qualified tiers (basic, advanced, and all qualified) of all product groups received more preferential placement than non-qualified models, although not all differences are statistically significant. Results were more mixed for sale pricing and discount. Qualified air cleaners, freezers, and room air conditioners were all more likely to be on sale than non-qualified models, but dryers, refrigerators, sound bars, and washers were less likely. Given that a model was on sale, qualified air cleaners, dryers, freezers, room

air conditioners, and washers received larger discounts than non-qualified models, but refrigerators and sound bars received smaller discounts. These results for preferential placement and sale quantity are consistent with the program impact inducing changes in promotional decisions by the retailers, although without pre-program data we are unable to compare relative changes.

Table 5-8. Summary of Retailer Behavior Analysis Using Shelf Survey Data

Product	Tier/Group	Sales Analysis	Placement	On Sale	Size of Discount	Incented by PG&E?
Air Cleaners	Qualified	--	0	+	++	
	Basic	--	0	0	++	Yes
	Advanced	--	0	0	++	Yes
Dryers	Qualified	++	++	-	++	
	Basic	++	++	-	++	Yes
	Advanced	++	+	--	+	Yes
Freezers	Qualified	++	0	++	++	
	Basic	++	0	++	++	Yes
	Advanced	--	0	0	0	Yes
Refrigerators	Qualified	0	++	--	--	
	Basic	0	0	0	--	No
	Advanced	++	++	--	--	Yes
Room ACs	Qualified	0	++	+	++	
	Basic	0	0	+	++	Yes
	Advanced	0	++	0	0	PY 2017 forward
Sound Bars	Qualified	++	+	--	--	
	Basic	--	+	--	--	PY 2016 only
	Advanced	++	0	0	0	Yes
Washers	Qualified	++	++	--	++	
	Basic	++	++	--	++	No
	Advanced	--	0	-	++	Yes

Key: "++" and "+" signify large and small increases, respectively, relative to non-qualified models. "--" and "-" signify large and small decreases, respectively, relative to non-qualified models. "0" signifies no statistically significant difference was found.

These results also help corroborate the sales increases calculated in Section 5.2, especially with respect to product placement within the store. Although the correlation is not perfect, the tiers with short-term sales increases tended to be those with the strongest evidence of receiving more preferential treatment in terms of placement within the store. The qualitative correspondence between these results is shown in Table 5-9.

Table 5-9. Qualitative Comparison of Sales Modeling Results and Retailer Behavior Analysis Using Shelf Survey Data

Product Category	Tier	Years Incented		Increase in Sales Above Baseline?	Preferential Treatment by Retailers?
		PY 2016	PY 2017		
Air Cleaners	Basic	Yes	Yes	No	Yes
	Advanced	Yes	Yes	No	Yes
Air Conditioners	Basic	Yes	Yes	Indeterminate	Yes
	Advanced	No	Yes	Too few sales	Yes
Dryers	Basic	Yes	Yes	Yes	Yes
	Advanced	Yes	Yes	Yes	Yes
Freezers	Basic	Yes	Yes	Yes	Yes
	Advanced	Yes	Yes	Too few sales	No
Refrigerators	Basic	No	No	No	No
	Advanced	No	Yes	Yes	Yes
Soundbars	Basic	Yes	No	No	Yes
	Advanced	Yes	Yes	Yes	No
Washers	Basic	No	No	Yes	Yes
	Advanced	No	Yes	No	Yes

INCREASE IN PROGRAM-QUALIFIED SHARE LEADS TO INCREASES IN MINIMUM REQUIREMENTS FOR ESRPP PRODUCT TIERS (M1.2)

As a result of increasing sales share of program-qualified models, the national ESRPP collaborative has "ratcheted up" tier eligibility requirements for several products—including air cleaners, refrigerators, and washers—in order to readjust for high market share.

Interview results show that PG&E has been an important contributor to these tier eligibility decisions. PG&E's subcontractors perform market analysis on product categories to determine if and when adjustments should be made. Interviewees noted that while this can be a difficult task, it is critical to the long-term success of the program.

External collaborator interviewees expressed differing views on whether ESRPP should continue to set qualification levels that are higher than the requirements for ENERGY STAR. One interviewee believed that this complicated the process and that retailers had trouble understanding it. Another interviewee believed that some sort of tiers are necessary to help drive the market forward.

ESRPP HAS SCALE TO INFLUENCE PARTICIPATING RETAILERS (M3)

Nationally, ESRPP has achieved coverage of an estimated 18% of the US population.[33] This is lower than ESRPP's stated goal of 30% (based on retailer input), but high enough to keep retailers engaged with the program.

Several interviewees had expected early on that other California IOUs would have joined the ESRPP program, thereby giving it substantially greater scale, and were surprised when this did not happen. One interviewee noted that retailers believe having additional IOUs join the program would constitute an important milestone for the program.

Interview results show that some market actors in the traditional retail supply chain already know and care about ESRPP, which according to program theory was not expected to occur for another five years or so. For instance, one interviewee (a representative from an electronics manufacturing company) already had exposure to ESRPP, largely through connections in utilities, mainly PG&E. The interviewee stated that ENERGY STAR—and thus ESRPP—moves the market in a way that is surprising for a voluntary specification.

ENERGY STAR SPECIFICATION CRITERIA FOR PRODUCT CATEGORIES BECOME MORE STRINGENT (M4)

Interviews conducted as part of this evaluation research suggest that ESRPP inputs have had a direct effect on the development of ENERGY STAR specification criteria, though not enough time has passed for these effects to be fully realized in the adoption of new specifications. Directional evidence for this assertion comes from interviews with external collaborators, who attest that the sales data and technical input from ESRPP has been a critical piece of the specification development process. Several data points illustrate this:

- PG&E provided data (from ESRPP) and information for the CA IOU Comment Letter on the Version 4.0 Discussion Document for sound bars and for the second round of comments on Version 4.0.
- PG&E submitted comments on the ENERGY STAR Discussion Guide for Air Cleaners. The Guide includes acknowledgement of ESRPP's involvement in opening the spec revision. PG&E additionally provided insights from web scraped data to EPA, showing that the market share for ENERGY STAR

[33] This figure is based on PG&E estimates.

models was higher than reported by the EPA, and helping make the case for a revision to the existing specification.

- External collaborators interviewed as part of this evaluation believe that the process by which PG&E works with the ENERGY STAR program to advance ENERGY STAR-qualifying requirements works well. Interviewees noted that the specification development process is typically spread out over a long time period, and so the types of interactions between PG&E ESRPP staff and ENERGY STAR will depend on the schedule for draft specification revisions or new specifications.
- External collaborators indicated that the full category sales data provided by retailers as part of the ESRPP program requirements was a notable achievement. EPA staff reported that this data has been an important piece in specification-setting efforts and has allowed them to drive specifications higher than they otherwise would have been able to do.
- External collaborators also noted that PG&E has been able to leverage its own internal resources to address technical testing needs related to two ESRPP products (soundbars and air cleaners) and that the results of this testing have also aided specification-setting efforts.
- As one external collaborator noted: "[Specification-setting] runs the risk of not having a balanced pool if utilities are not accounted for. ESRPP helps EPA defend against pushback from manufacturers and others."

5.4 LONG-TERM OUTCOMES

In the long-term (7-10 years), the PG&E ESRPP Program Pilot is expected to result in three key outcomes:

- Manufacturers should increase the number and type of efficient models in ESRPP-targeted product categories in order to address increased demand (L1).
- This should correspond to an increase in the national market share of efficient models (L2).
- Lastly, federal standards should increase sooner than they otherwise would have for ESRPP product categories (L3).

As the PG&E ESRPP Program Pilot has only been operating for less than three years, the evaluation team did not assess any of the long-term outcomes for this report (Figure 5-5).

Figure 5-5. Long-Term Outcomes

6. CONCLUSIONS AND RECOMMENDATIONS

The 2016-2018 PG&E ESRPP Program Pilot Evaluation research resulted in the following key conclusions and recommendations:

Conclusion 1: The PG&E ESRPP Program Pilot has implemented key activities necessary for the program to operate effectively, but impacts vary by product category. This reinforces the need for the program to have product category-specific strategies and goals that can be tracked and periodically reevaluated. It also suggests that not all product categories may be suitable to include in the ESRPP program.

> **Recommendation 1.1: Continue to develop product category-specific strategies and targets that are tailored to each product.** Additionally, for product categories where an increase in market share is not the primary objective, make sure that another objective has been identified and is clearly documented. For instance, there may be products where the primary objective is to help advance ENERGY STAR specifications. In these cases, there should be a specific need that ESRPP can address (for instance, by providing full-category sales data). To ensure that credit is given to PG&E, it is critical to document the impacts that these data have on subsequent developments for specifications, codes, or standards.
>
> **Recommendation 1.2: Product categories for which we have not yet observed an increase in sales or assortment share should be closely monitored to ensure they are making reasonable progress toward the objective for that product category.** For some product categories, the value of obtaining full category sales data from retailers may provide substantial benefit to PG&E efforts to advance specifications, codes, and standards. In these cases, there is an argument for keeping these product categories in the program, assuming that the relevant sales data can be used to advance voluntary or mandatory requirements (see Recommendation #1.1 above). It may be prudent to make downward adjustments to the incentive amounts for these product categories to reflect this strategy.

Conclusion 2: Analysis of sales data shows short-term increases in the sales share of program-qualified models for 5/11 product tiers, or 4/7 product categories currently targeted by PG&E ESRPP. At the same time, we observed preferential retailer promotional efforts and assortment increases for many of these same product categories. Collectively, this provides evidence that the core ESRPP program mechanism is working for these product categories/tiers. Our analysis indicates that the ESRPP intervention is linked to a statistically-significant increase in sales for dryers (basic and advanced), freezers (advanced), and soundbars (advanced). Additionally, we see a small but statistically-significant upward trend in the shelf assortment of program-qualified models on store shelves—a mid-term outcome which is expected

to follow increases in program-qualified sales. Collectively these findings provide supporting evidence that, for some product categories, the core ESRPP intervention is having some effect.

Conclusion 3: National ESRPP program processes could be improved by adopting a simplified approach for defining tiers within a product category and, to the extent possible, aligning these tiers with ENERGY STAR requirements. An important feature of the ESRPP program design is the ability to "ratchet up" tier requirements as program-qualified share increases for these product categories. To date, the ESRPP collaborative has used a flexible method in which tier eligibility requirements are aligned annually with ENERGY STAR specifications except in cases where the market share for that product category is already high. In these cases, the tier requirements have been set to "ENERGY STAR + XX%" (where the precise percentage varies based on the current program-qualified market share). This is a necessary adjustment for the program to make. However, in some cases it has caused logistical difficulties for the program and for retailers because it becomes more difficult to determine which models actually qualify for each tier.

> **Recommendation 3.1: In the future, PG&E should work with other program sponsors to explore simplifying the qualifying requirements used for the national ESRPP program and, to the extent possible, keeping these qualifying requirements aligned with ENERGY STAR definitions.** For instance, ESRPP could choose to align qualifying requirements with ENERGY STAR Most Efficient (ESME) in categories where there is such designation. In categories that lack ESME, there may be value in working with the EPA to establish such a designation.

Conclusion 4: The full category sales data provided by participating retailers are a valuable tool, particularly for facilitating the development of specifications, codes, and standards. Interviews with external collaborators indicate that these data have already been used to facilitate the development of ENERGY STAR specifications. Further research has revealed that these data do not exist anywhere outside of the ESRPP efforts, making it an even more valuable resource.

> **Recommendation 4.1:** Given the long-term program goals of changing mandatory and voluntary specifications, PG&E should continue to work with regulatory bodies to provide data and analysis to accelerate the adoption of these rules.

Conclusion 5: PG&E's ESRPP program pilot is highly influential within the national ESRPP collaborative effort. Interviews with external collaborators provide evidence that PG&E is considered by other program sponsors and

collaborating agencies to be one of two primary drivers of the national ESRPP effort, the other driver being the Northwest Energy Efficiency Alliance (NEEA). In particular, it appears that PG&E and NEEA are driving much of the codes and standards advocacy work.

Conclusion 6: As the PG&E ESRPP Program Pilot continues to operate moving forward, the current baseline approach (i.e., a pre/post model averaging baseline) will become less useful as the pre-period sales data become outdated. Therefore, it will become increasingly important to use a baseline approach that is able to account for new developments and external changes in the market.

> **Recommendation 6.1: Moving forward, the PG&E ESRPP Program Pilot should adopt a baseline approach similar to that employed by NEEA to help understand and assess market transformation effects due to the ESRPP program**. There are several benefits of using a baseline approach similar to that utilized by NEEA: (1) the approach has already been in use for some time, (2) it is transparent and flexible, and (3) using such an approach would facilitate evaluation consistency across two of the most important ESRPP program sponsors.