

200 Programs, One Evaluation Budget, One New Approach: A Recounting of the Great “Program Assessments”

Christina Torok, Evergreen Economics, Berkeley, CA

Michael Rovito, ERS, New York City, NY and Jennifer Fagan, Itron, Madison, WI

ABSTRACT

This paper presents the methods, outcomes, and lessons learned from implementing an innovative new study design in the field of EM&V. It shares perspectives on success in achieving its various objectives; where it was effective; and the areas in which it fell short and why. It suggests ways the study design could be improved. The study was implemented during the California 2010-2012 program cycle using approximately \$2M budget of EM&V funding allocated to the process evaluation of California nonresidential portfolio of programs. Implementation was unusual in that it was jointly funded and managed by the California PUC and the IOUs. Discussions of this paper will include considerations of the management structure in study outcomes.

The Program Assessments approach draws on techniques and findings developed for the California Best Practices Study¹. The best practices framework was modified to complete structured, comprehensive reviews of program features. The study also sought to update and refine current best practices. The purpose of the Program Assessment approach is to review the relationship of program design and implementation to the achievement of portfolio goals, within the context of markets, policy, economy and technology. The approach is grounded in contextual research: reviews of program design documents, previous evaluations, industry papers, economic and technical trends, and policy features. Reviews of quantitative data (budgets/tracking data) were also completed and incorporated, as data was available. The primary data collection effort was a large number of in-depth interviews conducted with program implementers, administrators and managers.

Introduction

This unique project arose from the evaluation challenges associated with California's large and diverse portfolio of nonresidential energy efficiency programs. During the 2010-2012 cycle California had over 200 primarily nonresidential programs across the four Independently Owned Utilities (IOUs). The size and complexity of this portfolio precluded the methods traditionally used to meet the basic need for due diligence review and feedback. Given the particulars of timelines and budgets, it was not feasible or cost effective to do an impact evaluation for each individual program or, even, a process evaluation.

In response to these evaluation challenges new methodologies were developed and implemented. These methods achieved extensive breadth of scope through the use of highly prescribed evaluation structures that focus on observable characteristics, that make full use of

¹ <http://eebestpractices.com/>

existing data / documents, and that leverage the knowledge and expertise of professionals closest to the programs, from implementation to evaluation and regulation.

The Program Assessment Reports were published in 2012 and early 2013, and they can be found on energydataweb.org. Findings from the Lower Rigor Assessments were published as part of the Custom Impact Interim Report published in December 2012 and also found on energydataweb.org.

This paper presents the techniques, key findings and lessons learned from implementing the Program Assessments technique over the 2010-2012 evaluation cycle. We relate key lessons learned from implementation of this method and identify the challenges and areas of success.

Purpose

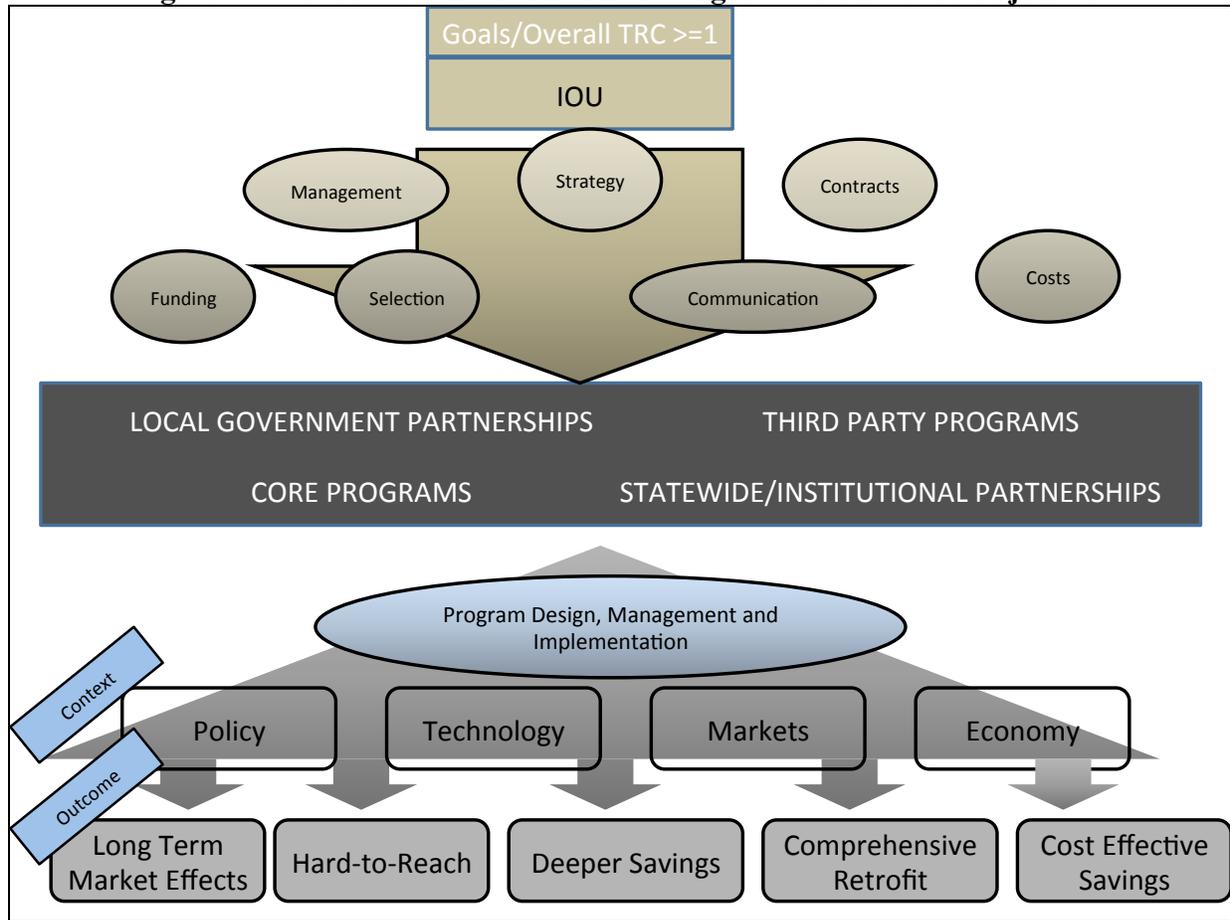
Two project methodologies were ‘road tested’ in the 2010-2012 program cycle to address the need for due diligence review and timely, actionable feedback across the California Nonresidential portfolio—“Program Assessments” and “Lower Rigor Assessments”.

The IOUs and the California Public Utilities Commission (CPUC) jointly funded and managed the Program Assessments projects as part of nonresidential process evaluation efforts. The primary purposes of the Program Assessments were: documenting program practices and investigating how well those practices achieve intended outcomes; providing lessons learned and recommendations to improve program strategies; and supporting resource allocation and decision making at the portfolio-level. While not an impact study, the method takes an integrated evaluation approach in performing the assessments, with consideration of elements related to both process and impact evaluation.

The study approach is wide ranging in its inputs and objective; it requires a documentation of the design of programs and an understanding of their target markets, technologies used, objective outcomes, and the underlying program theories. It examines qualitative program elements and characteristics (marketing, quality control, audits, technical assistance, innovation / adaptability) and factors in quantitative program achievements using savings, TRC, NTG, and other data from the impact evaluations. Moreover, it considers the context of program objectives – the particular features of the markets and/or technologies targeted by the program. It compares program implementation processes to known Best Practices, and it seeks to leverage findings to understand strengths and develop new Best Practices by “model” (e.g., core deemed vs. third-party calculated) and desired outcomes (long term market effects, comprehensive retrofit, hard-to-reach, cost effective savings).

Figure 1 below illustrates the portfolio construction, the major components of the portfolio and the portfolio goals. Program Assessments sought to understand the lower half of the diagram: the relationship of program design and implementation to the achievement of portfolio goals, within the context of markets, policy, economy and technology.

Figure 1: The California Portfolio and Program Assessments Objectives



Methods

The Program Assessments methodology leverages techniques and findings developed for the previous energy efficiency Best Practices Study (see www.eebestpractices.com). The previous best practices study developed and applied a comprehensive model that laid out a structural taxonomy for components shared by all (or most) programs. The model provides the foundation for elements that were reviewed and compared with outcomes, as well as benchmarked to ‘known’ best practices. The best practices framework offers a structure and a process for performing reviews across a full range of elements that make up a program: design, management, implementation and evaluation, described in more detail in the “Best Practices: Origin and Background” section below.

The goal of the approach is to cost effectively and systematically provide meaningful feedback and lessons learned associated with the large pool of nonresidential programs. The design’s core element was to perform due-diligence reviews of programs by ‘benchmarking’ to known best practices, and to identify lessons learned and new best practices from program experiences. A priori, the large and varied set of programs was a promising context for developing insights related to new program elements and their associated strengths or challenges.

All 200 plus programs in the California statewide PY2010-2012 energy efficiency portfolio were grouped into categories of similar design and orientation, and assessments were performed for each group.

Each group assessment used a common set of research questions and a similar set of research tools. However, some adjustments were made to the approach to best accommodate unique features or objectives of each group. The group assessments were based on interviews with program managers, reviews of program documentation and collateral, and in certain cases, early findings from ongoing EM&V studies, to provide a more comprehensive review. Group assessments provide a characterization of each program, a review against known best practices, and the development of new best practices. These assessments focus on the four “controllable” program components: design, management, implementation, and evaluation, per the Best Practices Program Decomposition Model.

This decomposition into components and sub-components provides the framework for a comprehensive and systematic approach to program review. Further it allows for targeted comparisons of program elements across programs, and well as to their intended outcomes, and to ‘known best practices’.

Best Practices: Origin and Background

The original Best Practices study was conducted in 2003-2005. The study resulted from a mandate by the California Public Utilities Commission to establish a database of programmatic excellence for use in developing new energy efficiency programs and improving existing ones. The term “Best Practice” refers to the business practice which, when compared to other business practices that are used to address a similar business process, produces superior results.

Based on the previous Best Practices work, this study used a *Program Decomposition Model* to break down and analyze each program and associated Group. Program decomposition refers to the process of disaggregating programs into their underlying subparts to allow for analysis of specific program features. The decomposition model provides a common framework for systematically evaluating and comparing program performance on a qualitative basis. The premise behind the original project, and this current extension of that work, was that a program is more likely to succeed, if it is known what others have done well and what mistakes are likely able to be avoided. Given, also, that field testing an idea is the only true way to understand how it will work and provide evidence that it does work, then by aggregating the knowledge gleaned from broad sets of similar programs one can help each individual program from reinventing the wheel.

Two levels of decomposition were done – a primary decomposition into components and a secondary decomposition into sub-components. Programs were disaggregated into four primary components for analysis: program design, program management, program implementation, and evaluation. In a second step, they were further broken down into sub-components as discussed below:

- ***Program Design.*** Program design subcomponents are focused on laying a solid foundation for a successful program.
 - ***Program Theory/Linkages & Partnerships.*** Good program design begins with good program theory and a complete understanding of the marketplace. Programs that demonstrate a clear “story” and understanding of the market, and have developed the

- right linkages and partnerships to successfully target that market, are likely to be more successful than programs that lack such characteristics.
- **Program Structure/Steps/Processes & Procedures.** Successful energy efficiency programs require well thought-out processes and procedures. Programs that clearly articulate the steps involved in implementation, as well as clearly delineate management responsibilities and structures have a higher likelihood of succeeding relative to those that do not.
 - **Program Management.** Program management is the command and control center that drives the implementation process.
 - **Project Management.** Effective project management represents the ability of the implementer to cost-effectively manage all aspects of the programmatic process by effectively executing the management/organizational plan. It is especially critical for implementers of large, complex programs or programs with multiple sub-contractors or other partners.
 - **Reporting & Tracking.** Tracking is defined as the systems and units of measure that provide an indication of program participation, budgets, markets and other program data. Choosing the right unit of measure to track a program can also be a predictor of success. Clear concise reports that track, for example, progress towards milestones and current expenses compared to projected levels are invaluable to program managers.
 - **Quality Control & Verification.** Systems for assessing the quality of program delivery, and for verifying the accuracy and prudence of tracking data, equipment and payments, are key to satisfied customers and successful programs.
 - **Program Implementation.**
 - **Outreach/Marketing/Advertising.** Good outreach, marketing and advertising efforts should result in relatively high program awareness, knowledge, and participation levels, all of which can be measured quantitatively. Success metrics, such as \$ per end user made aware or knowledgeable about a program can be used to benchmark and identify the most effective practices.
 - **Participation Process & Customer Service.** The ease or difficulty of a program's participation process, and the associated customer service support, can both be critically important indicators of ultimate program success. Some programs that may have all of the other attributes of success may be unsuccessful simply because the process of participation is unduly burdensome.
 - **Installation & Delivery.** Installation and Delivery picks up the implementation process at its finale and determines to what extent the program's implementation and design features carry through to installation of measures or adoption of behaviors.
 - **Program Evaluation.** Programs that are carefully evaluated and adjusted to ensure their effectiveness, and that can rapidly adapt to actual and changing market conditions, are more likely to be effective.

Lower Rigor Assessments: Origin and Background

As described above, there were two approaches that emerged in response to the need to cost effectively provide feedback across a large swath of programs in a short timeframe. In addition to Program Assessments, “Lower Rigor Assessments” were launched as part of the Custom Impact research. The assessments were a short structured review of custom projects, with the objective of providing qualitative feedback, not quantitative impact findings. One of the overarching objectives of the assessments was to determine the extent to which recommendations from the 2006-2008 Impact Evaluations had been implemented, and to determine if the implementation of those recommendations may have resulted in higher impact realization rates versus the previous program cycle. With these objectives in mind, the assessments focused on the procedures surrounding the determination and documentation of ex-ante savings, including baseline-related methods, procedures, estimates, and assumptions. The assessment itself was based on desk reviews of project application and program documents, and in some instances, on-site or telephone verification. A total of 300 additional sample points were drawn for the lower rigor assessments effort. The results of the lower rigor assessments were presented to program implementers and other stakeholders prior to the end of the program cycle and well before the final impact evaluation findings were complete. The California IOUs generally found the assessment reports useful as they provided the basis for critical and actionable feedback, and they did so before the results for program cycle were final. (These are available to the public in the Custom Impact Evaluation Interim Report, on energydataweb.org).

Implementing the Program Assessments Study

Implementing the structured methodology described above on a diverse and sizeable sample led to interesting challenges and successes. The following subsections address those issues with regard to how the sample was divided and with regard to how data was collected.

Program Groups and Assignments

One of the core challenges of implementing this study was the sample’s size and diversity. The programs targeted a variety of markets, had differing objectives, and leveraged diverse program models. Included were: the core programs, run by the utilities, with long-running offerings like the calculated (custom) programs, the deemed (rebate) programs, and the audit and pump test programs; nearly seventy local government partnership programs that worked with municipalities to tailor offerings to their facilities; more than a dozen statewide government and institutional partnership programs that develop deep relationships with large organizations that have big footprints like the California prison system or university system; and more than seventy five third-party programs, which spanned agricultural, commercial, and industrial sectors and were implemented by contracted entities, not the utilities.

Having one study team analyze the more than 200 programs would have been extraordinarily difficult. In order to facilitate a meaningful level of analysis, the programs were divided into groups with shared characteristics. The groups were organized as follows:

- Core Calculated
- Core Deemed
- Core Audit and Pump Test
- Local Government Partnerships

- Statewide Government and Institutional Partnerships
- Third-Party Commercial
- Third-Party Industrial and Agricultural

The decision to split or consolidate the main organizational groups was done mostly on the basis of size. For example, the core programs represent a significant portion of the total savings and budget of all the programs. Add to that the fact that their delivery models were significantly different and it became an easy decision to split them. Similarly, third-party programs were too numerous and divergent to be done all together. The decision was made to keep the industrial and agricultural together as they required similar non-standard process expertise to execute, whereas the commercial programs focused on more commonly available technologies.

The study team was made up of six different evaluation firms and the program groups were distributed among the teams. In certain cases teams collaborated on a single program group in order to match the relevant expertise of the teams with the needs of the program group. Overall study management was assigned to one firm. Weekly meetings were held to discuss shared challenges and study objectives. Templates for reports and interviews were distributed. It is important to note, though, that with this implementation model, the program group study teams were allowed some level of independence. The differences across program groups were sufficient enough to require, in some cases, major divergences from templates and overarching study practices. For example, partnerships were a very challenging group to fit into the model. These programs could be argued to be the most diverse set, and in addition diverge from traditional models substantially. Even program groups that conform to traditional models more closely needed tailoring for example, the best practices framework discussed above had to be adapted for third-party programs to include the component of utility oversight and contract management.

Data Collection Practices

The program assessments data collection strategy depended on collecting and integrating diverse sets of information. The idea was to draw both quantitative and qualitative data together to paint a more holistic picture than is possible in process or impact evaluations alone. Put another way, the team hoped to answer the what, how, and why all at once.

The core data collection technique was the long-form interview with program staff and related contributors. The interviews were based on the same template that aligned to the best practices decomposition model methodology, but were then adapted for each of the specific program groups. The hope was to develop interview guides that were sufficiently similar as to enable cross-program-group comparison, but sufficiently tailored to be useful for a specific program group. A master question tracker was developed so that shared questions could be edited collectively without inhibiting the necessary independence of each group. Certain pairings – for instance, the two third-party program groups – proved challenging since they required significant deviations from the core template, but also had to have significant blocks of shared questions across the two independently-studied program groups.

In addition to challenges of developing the interview guides, there were challenges in collecting the data from those interviews. The interviews were long-form and open-ended. An excel-based interview response tool was developed to help capture the information provided in a

systematic, searchable, and sortable way. This tool was ultimately critical for allowing responses to a single question to be easily compared across dozens of programs and multiple responses per program. One issue that arose during this process was the divergent ways in which interviewees interpreted questions. Getting the sentiment of an answer to align to the right question in order for easy comparison required critical analysis on the part of the scribes. And, as with all interview responses, the data was self-reported – it needed to be triangulated against other inputs.

The information gleaned from the interviews was supported with a variety of additional sources. Other program evaluations were ongoing or had been performed recently: impact evaluation was ongoing; process evaluations had occurred for key programs; a variety of market characterization studies had been prepared; and the historical record of all previous evaluations was at our disposal. These documents and data provided important context to the study. Timing did become an issue, though: the impact evaluation data did not become available until after the completion of most study activities. This would have provided an additional layer of important data. That said, from the quantitative perspective, tracking data offered a clear window into program practices across models. By matching patterns in the tracking data to interview responses and other evaluation reports, the teams were able to develop a multi-dimensional understanding of what was happening with each of the programs and, when trends persisted across many programs in a particular program group, those trends were grouped into given delivery model categories, market target categories, and the like..

Findings

The findings associated with the implementation of this study were diverse. Executing a unique and novel methodology brings with it experience and lessons learned regarding the method itself. At the same time, it exposes aspects of the subject – in this case the California Nonresidential portfolio – that were not previously identified in studies or well-understood. The following subsections address findings that align with both of these dimensions of research.

A Review of Methodological Successes and Challenges

The Program Assessments approach, as conceived, offered both successes and challenges. With regards to challenges, the study suffered from some critical missing links, and the study outcomes would have been improved by their inclusion. The first was access to complete program cycle accomplishments, if even in the form of raw tracking system data. The programs were evaluated mid-cycle, and some programs may have many or even most projects in development until very near the end of the cycle period. This reduced the team's ability to compare objective outcomes with subjective assessments of the process by which those outcomes were achieved. The Lower Rigor Assessments were able to fill this gap quite well for many quality control and management-related research questions. There are interesting and complex matters associated with the Lower Rigor exercise and its outcomes, which likely merit the attention of a separate, dedicated paper. Lower Rigor results were not fully integrated across all of the Program Assessments Reports, due to both timing issues and varying degrees of relevance to custom projects.

The study leveraged perspectives of those working very closely with program delivery, management and/or implementation. This provides detail and familiarity, but the ability to triangulate or verify perceptions by comparing views from multiple perspectives is necessary to form robust conclusions. Multiple perspectives were missing in key areas. Despite the substantial

insights the approach did yield, it could not always translate those into actionable recommendations due to a gap in access to the full range of perspectives on particularly contentious or complex issues. Ensuring a greater diversity of perspectives – implementer, customer, regulator, and others – will help triangulate observations on portfolio drivers of success, identify bias in the self-reported findings, and ground recommendations in a context that can at least acknowledge the concerns of all relevant stakeholders.

A benefit of the Program Assessments approach and its reliance on in-depth interviews and secondary sources is that it enabled the study to identify and report findings early and often. This provided for mid-cycle presentations and interim findings memos to support the bridge funding and portfolio application process. It was able to hone in quickly on the issues that were the most contentious and challenging, and to gather experiential accounting of how those issues manifest in day-to-day operations and achievements.

A significant aspect of the findings that stemmed from this methodology is that they tended towards the overarching rather than program-specific. The study had a unique ability to get a good snapshot of the overarching interactions, or “systems” that comprise the basic framework of the portfolio, and the ways they played out through the various participants and stakeholders. This is reflected in the following subsection, but is also addressed here because it should be understood as a direct extension of the study methodology. The study audience was made to understand that the California portfolio is significantly integrated; with programs responding to portfolio-level or model-level drivers and influences that shape their ability to achieve goals. This finding flows directly from the nature of the Program Assessments approach, which gathers data systematically across a range of programs and allows cross-cutting observation in a way that other study models do not.

The study was also successful in promoting conversation and debate on the shared elements of the portfolio framework that were contentious and most challenging, for at least one participant or stakeholder. Some such issues that surfaced during this study led some study participants to observe the EM&V platform might not be the most appropriate placement for the vetting or resolution of such issues; there are legislative and judicial processes in place to reconcile some of these. It was argued that the Program Assessments approach provided one more arena, among an ample set of arenas, to debate these issues. In response, this model offers advantages in that it forces participants to organize and align their input to the program decomposition model enhancing the context and providing for comparisons across programs and against known principals of best practices. Feedback grounded in a structured program review, if garnered from a sufficient variety of perspectives, offers an efficient way to identify key policy and program mechanisms and the particular objectives and outcomes they affect.

Reconsidering a “Program” for Assessment

When considering the assessment of California nonresidential programs the most important and salient characteristic is the way we define and think about individual programs in the context in which they are constructed and operated. The ‘programs’ are not stand-alone autonomous units that can be added, subtracted or changed individually, as one might traditionally think of a program. The California nonresidential portfolio is a fabric of internal and external resources and controls. Each ‘program’ is like a piece of that fabric made of strings that run through the whole portfolio. It cannot be removed in-tact, and wouldn’t be the same if placed in another position in the portfolio. Nowhere is this more tangible and evident than in the Local

Government Partnership sector, where ‘programs’ are really guidelines or agreements for a menu of cooperative activities promoting energy efficiency.

Ultimately, the full set of information, incentives and services that reach the nonresidential sector is the result of the interaction of IOUs, private companies, nonprofits, institutions, local governments, and individuals operating within a cross cutting framework of incentive structures, information flows, responsibilities and reporting requirements. The framework is set forth by the IOUs in their best effort to meet the needs of their own shareholders, ratepayers, other stakeholders, and the governing regulatory body. In the end, considering any one piece or nonresidential ‘program’ in isolation overlooks critical contextual parameters that define choices and resources. Any such consideration would misrepresent both accomplishments and shortcomings. This overarching construct and the findings that point to its defining importance changes the flavor of the original conception of the evaluation challenges the Program Assessments approach was conceived to address; from the need to assess many diverse ‘programs’ (by measuring them against best practices) to a need to address those cross-cutting threads that comprise the framework or fabric of the NR portfolio.

An excerpt from the Third Party Agriculture and Industrial Program Group Report illustrates the concept well:

“While the 3Ps (Third Parties) have a fair degree of independence in implementing their programs, it is important to understand the substantial role that the IOUs play in ensuring that their 3Ps deliver cost-effective, verifiable, robust savings. Assessment of IOU management of the 3Ps is as important, if not more important, than the evaluation of the 3Ps themselves. The IOUs are responsible for prominent and important functions that can determine the success or failure of the portfolio of 3P programs as well as individual programs.

- The IOUs plan the portfolio structure and program linkages in order to promote cooperation or competition among programs and integration with other IOU offerings.
- Through the procurement process, the IOUs must ensure that 3Ps have the proper skills and the proper contractual and compensatory motivation to achieve both primary and secondary portfolio objectives cost-effectively.
- During implementation, the IOUs’ oversight must balance the need for control of the process, which is essential for savings quality, with the need for 3P independence, which is essential for innovation.
- Finally, during evaluation, IOUs can encourage improvement in their portfolios by providing clear and consistent feedback on program performance.”

The salience of portfolio level issues created or corrected by IOU management and planning is evident from this passage. This observation was replicated across the study groups and should be central to any efforts to reassess and reapply this study framework in California again, or in other jurisdictions hoping to make use of this approach.

Identifying and Understanding Portfolio Tensions

Stakeholders each have their own “utility-function”, i.e., a unique set of risks, benefits and payback. Risks and benefits of one stakeholder are often at odds with those of others’ and this is often intentional and necessary in order to balance power and maximize overall benefits. The Program Assessments study brought into sharp relief the different and often competing objectives and interests of stakeholders. While it is not a surprise stakeholders have competing interests, it is important to understand where those interests are most strongly in conflict with the portfolio structure; how they are perceived and how they effect daily program practices.. The Program Assessments study documented the range of effects or responses of program implementers and managers as they operationalized the guidelines comprising the portfolio framework. The study method encouraged stakeholders to discuss practices, guidelines or market related issues they found critical to their success, as well as those presenting the greatest challenges. The issues that commonly arose as the study was implemented surrounded portfolio guidelines that were perceived as strong impediments by the interviewee / stakeholder. Often study implementers were able to document opposing viewpoints as well, if their set of interviewees was sufficiently diverse. As discussed above, securing sufficient diversity and documenting opposing perspectives is critical. For example, the Institutional Partnership Program Assessments study encountered major challenges effectively leveraging energy efficient program offerings due to different funding and planning cycles. From a policy or ratepayer perspective, making a long term grant or providing for a separate funding cycle accommodating various long planning horizons has implications in terms of risk and ratepayers control over energy efficiency commitments and spending levels. The study was not able to resolve this conflict, but it was able to report on the ways in which the conflict played out in the process of program implementation, daily practices and planning. In another assessment, for Third Party Commercial Programs, interviewees reported they were not always given a reasonable time to ‘ramp up’ their programs before being measured by performance. Third parties also report finding the uncertainty of cycle-to-cycle funding a significant business challenge. In many cases they articulated in detail the ways in which this aspect of portfolio design changed their business planning and program implementation. From an IOU and regulatory perspective, ‘relaxed early cycle performance standards’ and ‘early next-cycle funding decisions’ are in direct conflict with performance based provider selections.

These are two examples of how the study methodology does not yield solutions, but provides a high resolution ‘snap-shot’ of the portfolio design in operation. Moreover, the rigor of the structured and comprehensive program review process elevates the operational findings by forcing them to fit into very specific area of the ‘decomposition model’ and where elements can be compared across programs and against known principals of best practice..

Conclusion

The California nonresidential portfolio is so tightly bound together that for some issues the "program" can't be understood or interpreted without equal consideration for the guidelines that comprise the framework within which that program operates. This iteration of the Program Assessments study focused on measuring individual program performance against the best-practices measuring stick, but the Program Assessments method could be improved with more of a structural emphasis on understanding and documenting the governing framework. Relatedly, this shift should be matched with an effort to ensure a full compliment of perspectives are used

to inform reporting on how both the framework and the programmatic particulars are or aren't meeting challenges and attaining success. A benefit of the Program Assessments approach and its reliance on in-depth interviews and secondary sources is that it enabled the study to identify and report findings early and often. This provided for mid-cycle presentations and interim findings memos to support the bridge funding and portfolio application process

The most interesting findings from the 2010-2012 set of Program Assessments studies are related to understanding the incentives and drivers working across organizations and how they motivated the individuals delivering the portfolio. These finding also tied closely to the ways that program and portfolio frameworks reconcile competing objectives such as resource acquisition and market transformation. The insights offered through the Program Assessments approach help the audience better understand the mechanics of the trade-offs, and, as a corollary, how program or portfolio management might adjust the framework to achieve more or less of one element or another.

Accordingly, the Program Assessments approach is most useful in jurisdictions with large, complex portfolios that lend themselves to cross-cutting observations. For jurisdictions with few programs, the approach makes less sense to implement. Without the context of programs operating in parallel under similar rules, regulations, objectives, and budgets, it will be difficult to distill the effects of portfolio-level trends and drivers; to the contrary, large portfolios allow the evaluator to distinguish the portfolio or model-level signal from the program-specific noise.

References

Itron, Inc. 2012. "Program Assessments Study Work Plan, 2010-2012 Nonresidential Program Process Evaluation". Prepared for the California Public Utilities Commission and the California Investor Owned Utilities. www.energydataweb.com

Quantum Consulting, Inc. 2004 National Energy Efficiency Best Practices Study Best Practices Benchmarking for Energy Efficiency Programs. Prepared for California Best Practices Project Advisory Committee, PG&E. www.eebestpractices.com

Energy & Resource Solutions. 2012. "*Nonresidential Program Assessments Third Party Industrial and Agriculture Group*." Prepared for the California Public Utilities Commission and the California Investor Owned Utilities.

<http://www.energydataweb.com/cpucFiles/pdaDocs/859/3P%20Ag%20&%20Industrial%20Draft%20Report.pdf>

Morgan Marketing Partners. 2013. "*California Nonresidential Program Assessments Study IOU Core Deemed Program Group Report*." Prepared for the California Public Utilities Commission and the California Investor Owned Utilities.

<http://www.energydataweb.com/cpucFiles/pdaDocs/907/Deemed%20Final%20Report.pdf>

Itron and Energy & Resource Solutions. DRAFT 2013. "*California Nonresidential Program Assessments Study IOU Core Calculated Program Group Draft Report*." Prepared for the California Public Utilities Commission and the California Investor Owned Utilities.

Morgan Marketing Partners. 2013. “*California Nonresidential Program Assessments Study IOU Core Audit and Pump Test Program Group Report.*” Prepared for the California Public Utilities Commission and the California Investor Owned Utilities.

<http://www.energydataweb.com/cpucFiles/pdaDocs/905/NRA%20Final%20Report.pdf>

Navigant Consulting, Inc. 2013. “*Program Assessments Study Statewide Institutional IOU Energy Efficiency Partnership Programs.*” Prepared for the California Public Utilities Commission, California Investor Owned Utilities and Itron.

http://www.energydataweb.com/cpucFiles/pdaDocs/908/Statewide_Institutional_IOU_Energy_Efficiency_Partnership_Assessment%20Final%20Draft.pdf

Energy & Resource Solutions. 2012. “*Nonresidential Program Assessments Third Party Industrial and Agriculture Group.*” Prepared for the California Public Utilities Commission and the California Investor Owned Utilities.

<http://www.energydataweb.com/cpucFiles/pdaDocs/907/Deemed%20Final%20Report.pdf>

Evergreen Economics and Navigant Consulting. DRAFT. “*Program Assessments Study Local Government Partnership Programs Revised Draft Report.*” Prepared for the California Public Utilities Commission and California Investor Owned Utilities and Itron.

Itron and KEMA. 2013. “*2010-12 WO033 Custom Impact Evaluation Interim Report Final Draft*” Submitted to California Public Utilities Commission. www.energydataweb.com, Itron Work Order 36