Itron
Knowledge to Shape Your Future

Electric / Gas / Water
information collection, analysis and application

Maryland Strategic Evaluation Plan

Prepared for:

Crissy Godfrey
Demand Side Management Director
Maryland PSC
6 St. Paul Street
Baltimore, MD 21202

Prepared by:

Itron, Inc.
1111 Broadway, Suite 1800
Oakland, CA 94607
(510) 844-2800

August 19, 2010

Acknowledgements

This document was prepared by the independent evaluation team (Itron Inc.) with significant input and drafting contributions from the statewide utility evaluator (Navigant Consulting). Itron was hired by the Maryland Public Service Commission to review the evaluation work performed by the statewide utility evaluator and ultimately verify the energy impacts, peak demand impacts, and the cost effectiveness of the EmPOWER Maryland energy efficiency and demand response programs. Itron benefitted greatly from the review and assistance provided by the utility statewide evaluator in developing this document. Ultimately, however, the recommendations contained in this document are Itron's sole responsibility and Navigant retains the right to agree or disagree with the recommendations contained herein.

Navigant, with strategic input and oversight from Itron, was responsible for drafting Section 3 of this report, outlining the principles that it plans to use in making decisions on the scope and level of effort devoted to evaluating the savings impacts and effectiveness of the EmPOWER Maryland programs operated between 2009 and 2011. Itron has reviewed these recommendations and indicated within the document which of these Navigant recommendations we support. In many cases, Itron can not make a final decision on confidence intervals because there is not enough program specific information available yet. As a result, Itron will wait until we have the opportunity to review the final evaluation plan from Navigant before making decisions on confidence intervals for specific programs. Navigant also provided useful review and comments on other sections drafted by Itron including: Section 4 (Guidance on Contents of Utility Program Tracking Databases and Processes for Quality Control) and Section 6 (Proposed High Level Planning, Evaluation, and Reporting Schedule for 2010-2012). Finally we acknowledge the insightful review and comments provided by Crissy Godfrey (Director of Demand Side Management for the Maryland Public Service Commission).

Staff Acknowledgement

As project manager, I would also like to thank the many Itron staff members who contributed long hours to make these studies a success. I am particularly grateful to Joe Loper, Pamela Bagwell and Mike Rufo for their assistance in completing this evaluation plan.

Mike Messenger
Itron, Inc.

Table of Contents

Executive Summary

ES.1 Executive Summary

This strategic planning document was prepared by the independent evaluator team (led by Itron, Inc.) with significant input and review by PSC staff, the statewide evaluation team (led by Navigant, Inc.), the EmPOWER Maryland utilities, the Maryland Energy Administration and the Office of the People's Counsel.

The report will guide the statewide evaluator and utilities as they develop and execute their detailed evaluation plans for the 2010-11 programs. In addition to establishing a schedule for major milestones and deliverables, the strategic plan provides guidance to help with allocation of evaluation resources among different programs. In developing and finalizing this strategic plan, we hope to anticipate and resolve many high level evaluation issues in advance, identify differences in philosophy or approach now rather than waiting for them to emerge in final evaluation reports, facilitate timely process evaluations that can be used to improve program design and implementation, and strike a balance between best practices and getting the job done.

To be clear, this strategic plan does not eliminate the need for the independent evaluator team to scrutinize and review the statewide EE evaluation plans and the utility demand response (DR) evaluation plans when they are published later this year. Per its mandate, the statewide energy efficiency (EE) program evaluator will work with the utilities to design and conduct the EmPOWER Maryland and PJM EE program evaluations, including both impact and process evaluations. The independent evaluator will verify the results of the statewide EE evaluation conducted by the statewide evaluator, the utilities' DR evaluations, and estimates of program and portfolio cost-effectiveness, as well estimating co-benefits, and conducting other tasks as directed by the PSC.

ES.2 First-Year Priorities

The overall priorities of the EmPOWER Maryland evaluation in the first year are to (in order from highest to lowest):

1. Verify energy and peak savings from EE program portfolios for 2009-2010;
2. Estimate peak impacts from DR programs;
3. Calculate cost effectiveness of EmPOWER programs and portfolios;

4. Evaluate the effectiveness of program designs in recruiting customers and saving energy;

5. Develop plans to measure net savings; and

6. Begin to develop plans to address evaluation issues related to market transformation, multi-fuel impacts, remaining useful life of measures, dynamic pricing, etc.

The statewide evaluator's evaluation of the 2010 energy efficiency programs will focus on Priorities 1, 4 and 5 and be available in draft form in December of 2010 and final form in January of 2011. The independent evaluator's verification of 2009 &2010 programs savings estimates will be available in February of 2011. Priority issue 2, demand response programs, will be addressed in utility evaluations overseen by the independent evaluator. Final estimates of cost effectiveness through 2010, discussed as priority issue 3, will be provided in by the utilities in March 2011. Priority issue six – developing plans to address more complex evaluation issues related to market effects and multi fuel impacts – will be addressed at a very high level in these first reports and in more depth when the evaluation plans for 2011 programs are presented in April of 2011.

ES.3 Estimating Gross Impacts

We have established an aggressive schedule for the 2009-10 evaluation, requiring the utility reported impacts to be verified (and adjusted) by the statewide evaluator by January 2011. This schedule will be difficult to meet, especially given the slow ramp-up of programs and the late start of the evaluation process. To meet this schedule, the statewide evaluator will need to begin immediately to get measurement equipment and people into the field for priority programs while delaying the implementation of less urgent programs until the end of the summer and beyond.

Prioritization of programs will be reflected both in the timing of the evaluations and the level of rigor. The timing of the evaluations will depend on two factors. First, the statewide evaluator will evaluate programs according to how quickly the utilities expect them to ramp up. Second, estimating peak demand savings will require hours-of-use data for key measures to be obtained during the summer peak demand periods.

Evaluation rigor will necessarily vary for different programs depending on the current and projected share of overall energy and peak savings. At a minimum, all programs will be subject to desk review of the utilities' energy savings calculations. The desk reviews will be based on information provided in the Maryland Baseline Study, the recently completed Mid-Atlantic Technical Resource Manual (TRM) developed by the Northeast Energy Efficiency Partnership EMV Forum, and other relevant studies.

Based on review of the Maryland TRM, baseline study, and other relevant studies, the statewide and the independent evaluators may recommend changes to the ex ante savings estimates used by the utilities in their program tracking systems. These changes in ex ante savings assumptions will initially be based on the desktop review and will later need to be compared to any data collected in the field by the Navigant and Itron team as part of their evaluation of the 2009-2010 programs. These recommended changes in per measure ex ante savings values presumably will help explain any differences observed between reported and verified savings estimates at the program level. To date, there is no consensus on whether and when any recommended changes in ex ante values at the measure level from the statewide evaluation teams will be incorporated into the utilities' EMV annual reports, tracking systems and/or cost-effectiveness estimates. The timing for inclusion of these evaluation "findings" into key utility reporting documents, and the costs of doing so, are currently being reviewed by the Maryland PSC staff.

Based on discussion with the statewide evaluator, we have established minimum confidence and precision levels for various gross savings estimates. At the statewide *portfolio* level, energy savings estimates should have a 90% confidence that the estimated savings are within 10% of the actual values (i.e., 90/10 one-tailed test). High priority *programs* statewide should be evaluated at a 90/20 (one-tailed) confidence level. Lower *priority* program categories may be evaluated at the 80/20 (one-tailed) confidence. High priority programs are expected to account for approximately 80% of the total impacts.

To be clear, these confidence levels are for estimates of gross impacts; the confidence levels for estimates of net savings may be lower. Moreover, the independent evaluator may revise these target confidence levels, based on the relative priority of specific programs and the proposed emphasis on process compared to impact evaluations in the final evaluation plans.

A realization rate approach will be used to estimate impacts – i.e., estimates of savings will be based on the fraction of sampled utility-reported savings that can be verified. Recognizing that precision and confidence levels may be meaningless if the estimates are biased, the statewide evaluator will be careful to minimize various sources of bias, including non-representative sample design, measurement error, self-selection bias, model specification errors, program tracking system input errors, and issues with external validity. Discussion of potential bias and efforts to minimize bias will be included in the detail evaluation plans.

Questions of whether to develop estimates at the statewide program, utility service area, or utility program levels will be addressed in the detailed evaluation plans. The answers will depend on the specific objectives – e.g., assessing compliance with EmPOWER Maryland goals, estimating portfolio cost effectiveness, providing feedback to program implementers, etc.

The statewide evaluator will allocate evaluation resources to specific programs based on their overall contribution to program savings goals as part of detailed evaluation plans to be completed in August 2010.

ES.4 Process Evaluation

Due to the slow program ramp-up, in the first year evaluation we will place special emphasis on evaluation of program processes to provide formative guidance to program implementers. Process evaluations will include (but not be limited to) assessments of: staffing & trade ally experience and knowledge, consumer awareness of the programs and energy-saving opportunities, the cost effectiveness of key measures, customer motivators, and marketing approaches.

Process evaluations will be conducted for major program types and will include narrative descriptions of program theory as well as logic models. In future years, every effort will be made to complete process evaluations by June 1 to allow results to be reflected in the following year's program design.

ES.5 Evaluation of Demand Response (DR) Programs

The utilities are bidding available DR program peak demand reductions into the PJM base residual auction. PJM imposes requirements for these bids, including 90/10 (one-tailed) confidence and precision and schedules for verifying available peak demand reductions. The PJM may impose penalties if the bid capacity is not actually available and PJM reserves the right to require an independent audit of the bid reductions, to be paid for by the utilities.

Meanwhile, the independent evaluator is tasked with verifying the utility reported peak demand and energy savings for purposes of determining compliance with the EmPOWER Maryland peak savings goals and determining cost effectiveness of the DR programs. The utilities annual reports (submitted each year on January 30) should include DR resources available and called in the previous summer and the load impacts (and assumptions) projected for the coming summer. The Independent evaluator will then verify these estimates and report their findings back to each utility administrator and the Maryland PSC staff. It will be up to the administrator to decide if these values should then be changed before the final submittal to PJM before the summer season begins and or if additional evaluation work may be necessary to firm up the savings.

To facilitate this review, the utilities' 2010 DR evaluation plans should fully describe the methods and data, including discussion of key uncertainties and when and how ex ante values will be updated. The plans also should include projections of program enrollment,

estimated hourly load impacts for each event, and average hourly load impacts per participant.

ES.6 Estimating Cost Effectiveness

Given the priority the Maryland General Assembly has placed on measuring the cost effectiveness of the EmPOWER programs, Maryland Staff and the independent evaluator have worked with the utilities and other stakeholders to develop a practical and meaningful schedule and scope for cost-effectiveness estimates. Specifically, we recommend the following:

- The utilities should provide updated cost-effectiveness estimates for their major programs and portfolios by March 15, 2011. The cost-effectiveness estimates should be based on program participation and implementation through December 31, 2010 and the critical updates to ex ante assumptions and algorithms identified in the statewide evaluation of EmPower Maryland programs in January of 2011 and the independent verification of these savings estimates by Itron in February of 2011.
- Alternatively, if the Commission prefers, the utilities could provide preliminary cost-effectiveness estimates February 15, 2010 as part of their annual reports and then final estimates May 15, 2011 in time for their 2012-2014 program plans. The preliminary cost-effectiveness estimates would be based on ex ante assumptions used in the utilities' 2009-11 program plans, but would include updated participation and installation counts. The final estimates would be based on the full panoply of findings from the statewide and independent evaluators for the 2009-10 program activities.
- At least two cost-effectiveness estimates should be provided: 1) estimates based on 2009-2010 actual participation rates for each program, and 2) estimates for 2009-2011 programs, which will be based on 2009-10 actual participation rates and 2011 projected participation rates.
- Utilities should develop updated forecasts of program participation for 2010-2011 to be used in this cost-effectiveness analysis. The independent and statewide evaluators will review these projections to assess whether they are reasonable and sufficiently documented.
- During the course of their desk review of savings estimates, the statewide evaluator should review and recommend revisions to the ex ante savings assumptions and algorithms in the existing cost-effectiveness tools and/or tracking systems based on their review of the Maryland TRM, findings from the MD baseline study and other relevant studies. These recommendations should be shared and discussed with the

utilities and the independent evaluator on or before January 10, 2011 to allow time for integration into the final evaluation.

- The independent evaluator should review both the findings from the statewide evaluator on measure level savings estimates and the program-level methodologies used by each utility to develop parameters for the cost-effectiveness analysis; including incentive and administrative cost allocation, bundling of measures into programs, cost allocation for measures that save electricity or natural gas, discount rates, and common assumptions. Based on this review, the independent evaluator should recommend changes to savings assumptions or methodologies used to calculate cost-effectiveness estimates in its first verification report due on January 10, 2011.
- Utilities, with cooperation from their own internal contractors and/or the statewide evaluator, should provide all data necessary to assess cost effectiveness -- including collecting data participant measure costs and remaining useful lives of measures -- that will be needed for their update to the cost-effectiveness analysis.
- Realization rates used in the cost-effectiveness analysis should be developed based on program activity through September 30, 2010 and used in the initial estimates of portfolio cost effectiveness for 2009 and 2010 programs. These same realization rates should be applied to actual program activity through December 31, 2010 to produce annual estimates of program savings for 2010.
- Program and portfolio level cost effectiveness should be reported for each utility service area. The independent evaluator will be responsible for rolling up the results from all the utilities into a statewide total if this is requested by the Maryland PSC.

ES.7 Estimation of Net Savings

The EmPOWER Maryland Act and subsequent agreements have not stated whether the EmPOWER Maryland goals are to be based on gross or net impacts.[1] Regardless of whether the goals are based on net or gross impacts, estimates of net impacts will be needed for cost-effectiveness estimates and to support program design and implementation.

The statewide evaluator will use surveys and other methods to develop net-to-gross (NTG) ratios for all of the sampled projects. The statewide evaluator will rely primarily on survey

[1] Net impacts are gross impacts adjusted for free riders, spillover and rebound effects. Net impacts may be more or less than gross impacts. Free riders are participants that would have implemented the measures without the program. Spill-over is participants who implement additional measures as a result of participation in the program, or non-participants who implement measures as a result of the program. Rebound is when the reduced cost of energy services resulting from the efficiency improvements results in increased consumption of the energy services.

methods, but multiple methods will be used where appropriate to triangulate results. Likewise, multiple perspectives (e.g., consumers, retailers, manufacturers) will also be obtained. As noted above, confidence and precision levels will necessarily be lower for net savings estimates than for gross savings estimates, and in some cases precision and confidence targets may not even be meaningful for net savings estimates.

The utilities and the statewide evaluator should consider two-stage evaluation for some programs, including in particular large industrial and commercial custom programs. By interviewing program applicants early in the project implementation, more reliable estimates of free ridership can be obtained. Moreover, two-stage evaluation provides real-time feedback to program implementers about changes they could make to program eligibility requirements and marketing that would limit free rider participation. A two-stage evaluation process will need to overcome the concerns of some utilities that customers will be annoyed if they are denied rebates because "they were going to do the project anyway." Also, the statewide evaluator's scope does not include prospective evaluation of projects and measures, thus cost could be an issue.

ES.8 Deferred Evaluation Issues

The EmPOWER Maryland programs, targets and evaluations are ambitious and require a significant commitment of staff and contractor resources. Consequently, we have deferred resolution of some key evaluation issues for a later date, including: fuel switching impacts, multi-fuel demand impacts, expected versus remaining useful lives, social marketing effects, market transformation effects, dynamic pricing and the use of dual baselines (e.g., building codes versus standard practice). Although we do not expect these issues to be fully resolved, they should be discussed in the statewide evaluation plans.

ES.9 Program Tracking Systems

Evaluations are only as good as the tracking systems that support them. Development of reliable, accurate and user-friendly tracking systems is critical to the success of the EmPOWER Maryland evaluations, as well as the success of the programs. Achieving high levels of confidence and precision in the statewide evaluation will depend in large part on the accuracy of the tracking systems.

The strategic plan provides many recommendations related to the utility tracking systems, including (for example), that the tracking systems include a relational database with links between tables, standard field names, and unique record identifiers that remain with participants and measure types throughout the program. Assumptions and formulas should be fully documented and contain links to the relevant studies or calculations.

Robust quality control processes are vital. Manual transfer of data between databases and calculators should be avoided and systems should be in place for checking all data prior to entry into the tracking system. Automatic validation of fields should also be instituted where feasible to ensure the data is complete and accurate.

Ideally, each utility should combine program tracking systems for different programs into a *single* program tracking system. Over the long run, we urge the EMPOWER Maryland utilities to adopt the same tracking system to facilitate review and updating of assumptions and algorithms and to create uniform linkages with cost effectiveness and co-benefits calculators.

ES.10 Reporting

Clear, comprehensive and concise reporting of evaluation results is necessary to minimize the misinterpretation of evaluation results and provide context for the various estimates. The strategic plan recommends standard terminology to describe various stages of the statewide evaluation, namely that utility "estimated" savings will be called "forecast" savings, "actual" savings will be called "reported" savings, and savings estimated by the statewide evaluator (or the independent evaluator in the case of DR programs) will be called "verified" savings.

The utility EM&V annual reports should include: realization rates and explanations, discussion of uncertainty, best-estimates of program and portfolio cost effectiveness, recommended program design changes, full documentation of methods and assumptions, proposed modifications to the future evaluation design, and proposed future changes to the tracking systems, including changes in *ex ante* assumptions.

ES.11 Schedule and Milestones

Several key schedule constraints are driving the evaluation schedule, including

- January 30 – Annual Programmatic Savings Report from Maryland utility administrators to the PSC
- March 1 – Annual Report of the Maryland PSC to the Legislature on EmPOWER MD Programs
- May 15 – Post-installation EM&V report due to PJM

Other constraints include delays in program ramp up, tracking system development, and the PJM's requirements for evaluation of DR resources bid into the base residual auction.

The table below includes key evaluation deliverables and due dates agreed to by all stakeholders.

Table ES-1: Evaluation Reporting and Planning

Key Dates	Deliverables
June 14, 2010	Draft Strategic Evaluation Plan
Aug 15, 2010	Final Evaluation Plans (Navigant will stagger high and low priority programs)
Dec 1, 2010 and Jan 15, 2011	Draft and Final Statewide Evaluation Report of 2009-2010 Program Savings (Navigant)[2]
Jan 10, 2011 and Feb 15, 2011	Draft and Final Verification of 2009 and 2010 Statewide Program Savings Report (Itron)
Jan 30 annually	Utility Programmatic Savings Reports - Includes reported and verified savings for previous program year (e.g., 2010 programs for the Jan 30, 2011 report)
March 1, 2011	EmPOWER progress report to State General Assembly
March 15, 2011	Utilities submit cost-effectiveness analysis for major programs and the entire portfolio for program years 2009, 2010, and 2011 program years.
May 1, 2011	Post-Installation Report to PJM for program savings bid into the market[3]
May 1 annually	Final Process Evaluation Results and Recommended Design Changes – interim results to be provided to utilities throughout the year

[2] Note the January 15 estimate of verified savings for 2010 programs does not need to be filed with the Maryland PSC but should be completed and sent to each administrator and Itron by January 15 to allow each administrator sufficient time to include these savings estimates in their January 30 Annual Program Savings Reports. This will also give Itron sufficient time to complete its independent verification of these savings numbers by February 15 of each year.

[3] The May 1 date for a post-installation M&V report to PJM for 2010 is dependent on whether utilities choose to bid into the 2011/2012 PJM Capacity Market.

1

Introduction - Purpose and Intended Audience

1.1 Purpose of the Plan

This strategic evaluation plan was prepared by Itron Inc., in consultation with the Maryland Public Service Commission (PSC). Itron is working for the Maryland PSC as the Commission's independent evaluator of the Maryland investor-owned utilities' EmPOWER Maryland programs. Itron works with the PSC to provide guidance for and oversight of the Maryland utilities' program evaluations. Navigant is working for the Maryland utilities as the statewide evaluator to help design and conduct the EmPOWER program evaluations.

The purpose of this strategic evaluation plan is to:

- Provide guidance for use by the Maryland utilities and their Statewide Evaluator in developing Statewide Evaluation Plans for 2010-11 energy efficiency and demand response programs.[1]
- Propose criteria for use in deciding the scope of programs to be evaluated and the allocation of evaluation funds.
- Provide a detailed evaluation and reporting schedule identifying how and when evaluation results will be produced, reviewed, and ultimately used by the Maryland PSC to assess the overall effectiveness of the EmPOWER Maryland programs.

Importantly, this strategic evaluation plan is not intended to replace subsequent review of the statewide evaluation plans, but rather to anticipate and resolve high level issues in advance. Itron will provide a thorough and detailed review of Navigant's utility evaluation plans once they are completed in August 2010.

In addition, due to schedule constraints, this first Maryland PSC Strategic Evaluation Plan focuses only on the most important high level issues that must be addressed immediately. These

[1] There are likely to be two separate evaluation plans filed by July 1, 2010, one from Navigant for Energy Efficiency Programs of all five EmPOWER MD utilities and another separate plan from the utility program administrators (or their consultants) on their evaluation plans to estimate and verify the peak savings impacts of their demand response programs.

are primarily related to evaluation priorities, schedules, tracking system needs, reporting requirements, and reporting schedules. Subsequent updates to this Strategic Plan will provide direction and requirement for other important and on-going activities such as cost effectiveness procedures and refinement of other evaluation policy rules (e.g., for related activities such as baseline energy use specification criteria, multi-fuel impacts, fuel switching, and distributed generation).

The intended audience for this plan is the five utility program administrators, their evaluation and implementation consultants, the Maryland Energy Administration, the Maryland PSC staff and Commissioners, and all interested Maryland stakeholders who intend to review evaluation products and contribute their expertise.

1.1.1 Key Research Questions

Table 1-1 presents our understanding of the types of research questions that the various target audiences in Maryland would like to have answered by the program evaluations conducted over the next two years. This table was constructed based on review of the program legislation and discussions with program administrators and evaluation staff in Maryland. Given the budget for evaluation is constrained in Maryland to roughly two percent of expenditures, not all of these questions can be answered.

These questions are not intended to represent a comprehensive list, but rather to help illustrate the broad range of expectations that should be considered in developing evaluation priorities. These priorities will be expressed in both this strategic evaluation plan and the actual program evaluation plans to be completed by Navigant and the program administrators in July and August of this year.

Understanding the evaluation priorities and needs of each of these audiences will be important in helping Itron and Navigant allocate evaluation resources among different types of evaluations (i.e., impact versus process), different programs, and different levels of granularity. Tradeoffs will need to be made in deciding on the appropriate allocation of evaluation resources used to estimate load impacts at the statewide level, at the utility service area level, and at the individual program level for each utility. As the table above suggests, different audiences are likely to have different priorities for what questions are the most important to be evaluated and at what level they should be answered (statewide, utility or program). For example, the Maryland State Assembly might be satisfied with an evaluation of savings impacts at the statewide level to assure the goals were met while each utility may be more interested in receiving process and impact results for their specific programs.

Given these potential differences, we expect that different stakeholders in Maryland will have different perspectives on the evaluation priorities recommended by Itron in Section 2 of this

report. Itron and Navigant will attempt to balance these competing priorities in the final evaluation plans for energy efficiency programs to be completed in August 2010.

Table 1-1 Illustrative List of Research questions by Target Audience

Target Audience	Impacts - What happened as a result of the programs?	Process - Why did it happen & how can we do better?
Maryland Legislature	■ Did the sum of all EmPOWER MD utility programs reach the legislative goal of reducing per capita electricity consumption by 5% by 2011[2] and 15% by 2015? ■ Were the programs implemented in such a way that the benefits exceeded the costs - i.e., were they cost effective? ■ Did the programs produce actual reductions in the energy consumption of participating customers?	■ Which programs were the most successful in saving energy and why? ■ Which programs or portfolios of programs were the most cost effective and why? ■ If the programs are not on track to meet the goals, what changes are needed to ensure they will be met by 2015?
Maryland PSC	■ What level of program energy savings was verified for each utility portfolio by the independent evaluator team and how does this compare both to their forecasts of program savings in 2008 and the level of savings reported in their quarterly and annual reports to the PSC? ■ What assumptions related to energy savings estimates made in the planning process need to be revised before the next EmPOWER Maryland planning process? What were the largest uncertainties identified by the evaluators in current program savings estimates and other parameters related to the cost effectiveness of programs?	■ Why were estimates of savings at the forecast and reporting levels different than the savings levels verified by the independent evaluation teams? ■ Given the observed differences between forecasted, reported and verified energy savings, what can be done to improve the program planning and reporting processes? ■ What sectors or program areas hold the most potential for achieving additional energy savings in the next planning cycle?
EmPOWER Maryland Utilities	■ What were the factors that led to differences between forecasted program savings (based on customer participation forecasts) and reported savings (based on actual levels of customer participation)? How did these differences affect payments received in the PJM forward capacity market? ■ What factors led to differences between reported savings and verified savings for both energy efficiency programs and demand response programs? ■ Which programs performed the best and which programs need to be modified or considered for elimination based on the results from the statewide and independent evaluators?	■ What explains the differences between forecast, reported and verified savings and to what extent were these factors within the utilities' control? ■ What can utilities do to increase the cost effectiveness of their programs while still meeting the EmPOWER Maryland goals?
Maryland Energy Administration (MEA)	■ What EmPOWER Maryland programs benefited directly or indirectly from the support or existence of programs administered by the MEA? Can these effects be quantified?	■ What changes if any can the MEA make to its programs to maximize the chances that Maryland utilities meet the EmPOWER Maryland savings goals?

[2] Note, the MD utilities are only responsible for reaching 2/3 of the EmPOWER Energy savings goal by 2015.

1.2 Status of the Rollout of EmPOWER Maryland Energy Efficiency Programs (as of January 1, 2010)

In developing priorities for the statewide evaluation, it is important to also consider the status of the EmPOWER Maryland programs, some of which were not launched until early spring of 2010. In particular we need to understand progress each utility is making with respect to achieving its planned savings and expenditure goals. This will guide where evaluation resources might be most effectively spent in the short run.

Table 1-2 compares actual (reported) energy and peak savings to the estimated (forecasted) savings provided in the utilities' 2008 EmPOWER program plan filings to the PSC.[3] This table suggests that most of the EmPOWER Maryland utilities (with the possible exception of BGE) are experiencing some difficulty in rolling out and ramping up their programs due to either late start dates, slow ramp up rates (perhaps due to delays in hiring the needed staff and contractual support), or less effective than planned program designs. This table suggests evaluators should focus in the very short term on performing process evaluations for two reasons:

- Money spent on increasing program participation is likely to be more valuable than impact evaluations in the short run because increasing participation is a higher priority right now than performing a load impact analysis of some of the smaller programs, however it may be appropriate to perform a vigorous desk review of the ex ante savings estimates because faulty estimates could be part of the reason for low program participation
- Money spent on load impact evaluations could be partially wasted if there are not a sufficient number of participating customers to draw meaningful samples or develop any statistically significant estimates of program savings.

Thus we recommend a near-term focus on formative review of program designs and marketing strategies to help increase program participation. This of course needs to be balanced with the need to perform load impact analysis at least at the program level to help verify overall program accomplishments and cost effectiveness. Navigant will propose a plan to balance among these competing evaluation needs in its evaluation plans due in August of 2010.

[3] The utility annual reports use the terms" estimated savings" to refer to forecasts of savings made for these programs in 2008 and actual savings to refer to the program savings they reported in their annual report. Actual or reported savings represent the multiplication of the number of participants in the 2009 programs and the ex ante estimates of savings per customer or energy efficiency measure. In this report, we recommend the use of the term "forecasted" savings to refer to savings labeled by utilities as "estimated" savings and reported savings to refer to the program savings utilities have labeled as "actual" savings in their annual report. For convenience we use the utility definitions in this section to ensure readers can compare them to the figures cited in their reports but for the rest to the report we use our preferred terms, forecasted and reported savings.

Table 1-2 2009 Estimated vs. Actual Impacts by Utility

Utility	GWh/yr (2009)			MW (2009)		
	Estimated	Actual	% of Estimated	Estimated	Actual	% of Estimated
Allegheny	6.1	0.1	**2%**	3.1	0.0	**1%**
BGE	278.1	111.2	**40%**	49.9	14.8	**27%**
DPL	31.7	8.5	**27%**	7.0	0.6	**8%**
PEPCO	136.1	49.0	**36%**	28.0	3.4	**12%**
SMECO	20.1	0.2	**1%**	3.7	0.1	**2%**
Total	472.1	169.0	**34%**	91.7	18.9	**19%**

Sources: "Estimated" impacts from PSC staff email June 23, 2010. "Actual" impacts from utility EmPOWER Maryland 2009 annual reports. Per revised nomenclature suggested in Chapter 6, "estimated" is equivalent to "forecast" and "actual" is equivalent to "reported."

Table 1-3 compares estimated and actual program expenditures, revealing that utilities' actual program expenditures were much lower than the estimated program expenditures. Actual expenditures as a fraction of estimated expenditures ranged from 21% to 63% across the five utilities. This data confirm that reported expenditures are significantly lower than projections in the 2008 program plans; in large part because paid-out incentive costs and the associated savings are lower than forecasts. Review of the reported program participation counts suggests that most programs had not achieved the minimum number of participating customers needed to accurately evaluate program savings at the program level by January 2010. Exceptions to this rule appear to be the residential lighting programs and some other residential appliance and audit programs.

Table 1-3 2009 Estimated vs. Actual Programs Expenditures

Utility	Program Expenditure ($millions)		
	Estimated	Actual	% of Estimated
Allegheny	4.9	1.1	22%
BGE	42.9	27.1	63%
DPL	5.7	1.2	21%
PEPCO	14.0	3.3	24%
SMECO	4.7	0.6	13%
Total	72.3	33.3	46%

Sources: "Estimated" expenditures from PSC staff email June 23, 2010. "Actual" expenditures from utility EmPOWER Maryland 2009 annual reports. Per revised nomenclature suggested in Chapter 6, "estimated" is equivalent to "forecast" and "actual" is equivalent to "reported."

Customer participation for some of the major residential and commercial programs has picked up in the first quarter of 2010, but there are still a number of programs with an insufficient number of customers to warrant load impact studies. In developing the utility evaluation plans, Navigant will look at these participation trends on a program by program basis and make recommendations accordingly.

1.3 Preview of the Report Contents

Section 2 provides overall recommendations on six specific evaluation objectives that should be given the highest priority in the first two years of evaluations conducted in Maryland. Section 3 provides guidance on the expected levels of precision and confidence to be used in developing energy and peak savings estimates at the statewide, utility and program levels. It also includes the principles that will be used in developing the recommended mix of load impact and process evaluations for the first year. Section 4 provides an assessment of the existing program tracking systems and recommendations on how to improve their functionality for both program administrators and evaluators. Section 5 provides guidance on the contents of evaluation reports. Finally, Section 6 proposes a schedule for reports and other evaluation activities and Section 7 provides a preview of next steps in the evaluation planning process.

.

2

Guidance on Evaluation Priorities and Objectives for 2010-2012

This section provides guidance on the evaluation expectations and priorities for program evaluations and subsequent reporting between July 1, 2010 and June 30, 2012. These priorities have been reviewed and accepted by the Maryland Public Service Commission staff and represent their current understanding of the policy priorities of their Commission. These priorities are expected to guide the scope and relative level of evaluation effort proposed for:

- The evaluation of utility energy efficiency programs deployed from 2009 to 2011 (plans to be provided by Navigant), and
- Each utility administrator's evaluation plans for their demand response programs deployed from 2009 through 2011 (plans to be provided by each utility).

2.1 Summary of Evaluation Objectives

There are six evaluation objectives (in priority order) that should guide the development of the statewide evaluations plans from both Navigant and Itron. They are summarized below and discussed in more detail in the following sections:

- Section 2.2 provides verified program energy and peak savings associated with the energy efficiency program portfolios operated by the Maryland Utilities.
- Section 2.3 provided effective process evaluation results designed to increase the level of customer participation in, and net energy savings achieved by, new EE programs.
- Section 2.4 provided estimates of historically achieved peak-load impacts from demand response programs, and forecasts of future impacts.
- Section 2.5 calculates cost-effectiveness of EmPOWER Maryland program portfolios.
- Section 2.6 develops a strategy to begin to measure the net energy savings attributable to energy efficiency programs.
- Section 2.7 lays the groundwork for subsequent evaluation efforts related to fuel substitution, multi-fuel impacts, remaining useful life of in situ equipment, social

marketing programs intended to induce behavioral changes, and the measurement of long-term market effects induced by programs.

2.2 Provide Verified Program Energy and Peak Savings Associated with the Energy Efficiency Program Portfolios Operated by the Maryland Utilities

Determining utility compliance with the EmPOWER Maryland goals is a top priority of the statewide evaluation. The statewide evaluator shall produce an overall statewide evaluation plan in August 2010 and a preliminary estimate of verified energy and peak savings from the 2009 and 2010 energy efficiency programs by December 1, 2010.[4] The statewide evaluation plan shall specify the evaluation methods and sampling processes that will be used to verify the energy and peak savings from, at a minimum, the key energy efficiency programs in 2009, 2010 and 2011.

The statewide evaluation plan should also provide specific information on what types of information will be available to support the filing of formal 2012-14 energy efficiency program plans, which are due September 2011. This information should be provided to the utilities no later than June 1, 2011 in time to inform the draft 2012-14 program plans and discussions. In general, process studies and market research or assessments conducted by the statewide evaluator should be targeted for completion or provide interim process results in May of each year to provide administrators with sufficient time to consider modifications to their programs on an annual basis. This recommendation is discussed in more detail in Section 6 of this report.

The statewide evaluator also has the responsibility to support the development of energy efficiency program capacity bids made by the EmPOWER Maryland utilities in the PJM forward capacity market. To support this objective, the statewide evaluation plan should consider what information and results will need to be available by May 1 of each year to support the annual filings of the post-installation EM&V reports to the PJM. The intent of the post-installation reports is to confirm that the level of peak savings resources previously bid into the PJM market have been verified for the coming summer season. The post-installation reports are due at least 15 days before the start of the summer season on June 1 and thus it will be prudent to provide all necessary information to each utility by May 1, 2011 and annually thereafter. Utility administrators should also compare the amount of MW bid into each summer's PJM auction with the amount verified by the independent evaluator (Itron) in

[4] The details and rationale used to develop these dates is provided in Section 6 (Evaluation Planning Schedule).

its annual reports, which will be provided by February 15 of each year. Again, Section 6 of this report provides more details.

We recommend that the programs that are found to have low ratios of actual to estimated customer participation levels or savings based on a review of the latest program savings estimates from the quarterly reports should be strong candidates for Navigant to begin process evaluations in the second half of 2010. These process evaluations should include a review of the program's theory, development of logic models, and assessment of whether each program has sufficient staffing resources, assessment of current market awareness and adoption of key measures, assessment of participant cost effectiveness of key measures, and assessment of marketing approaches.

The relatively low levels of program participation in 2009 and the first quarter of 2010 make it even more important for evaluators to gain an understanding of which programs are likely to achieve significant levels of participation and savings by the end of this year. Without an updated forecast of likely 2010 program savings from utilities, it is difficult to make informed decisions about which programs to evaluate, how many customers to sample for each program, and at what level of depth. We strongly urge utility program administrators to produce a forecast of participation for year-end 2010 and share these with Navigant immediately.

2.3 Provide Timely Process Evaluation Results to Increase the Level of Customer Participation and Impacts from New Energy Efficiency Programs

In the early stages of any program launch cycle, it is critical to provide early feedback on the effectiveness of program designs and marketing in attracting participation and increasing the energy savings yield per unit of program expenditures. The statewide evaluation plan should balance the need to provide formative advice to program administrators with the need to verify the level of portfolio savings being achieved as described in objective 2.2 above.

Process evaluations are used to document program operations and provide information necessary to improve program operations and cost effectiveness. An important element of the 2010 statewide evaluation will be to inform program design. While gathering information to improve the performance of all programs, emphasis will be placed on underperforming programs and effective approaches that can be transferred to other programs. The process evaluation will provide immediate guidance to program implementers, as well as guidance to ensure best-in-class programs are filed by the utilities for the 2012-2015 program cycle due in September 2011.

Specific research objectives and tasks for the first-year process evaluation will include:

- Understanding and clarification of program goals,
- Review of data tracking system and recommendations for improving process of documenting program impacts,
- Exploration of statewide integration issues, such as confusion over multiple program designs in the market,
- Program marketing and outreach efforts, and
- Recommendations to help encourage program participation (this will include barrier research and target marketing efforts).

Before finalizing specific process evaluation objectives or tasks, the statewide evaluator will develop program evaluation plans for review and discussion with the Maryland EmPOWER utilities, the independent evaluator, and other stakeholders.

2.3.1 Program Theory and Logic Model

The statewide evaluator should work with utility program implementers to develop program theories and logic models for each program selected for a process evaluation. A program theory provides the underlying rationale for the cause and effect relationships of a program. A logic model is the graphical representation of a program theory, showing the flow between the activities and outputs, and short-term, intermediate, and long-term outcomes.

The design of the process evaluations undertaken in Maryland should be informed by the underlying theory of why program activities are expected to create specified outputs and outcomes. A theory-driven approach is particularly useful when a new program approach is being undertaken since many of the cause and effect relationships might be untested and implementation problems are likely to be more numerous.

Program theory and logic models also can be used to develop performance indicators that can be monitored over time and reported. Regularly providing such information to program managers can allow them to make mid-course correction in the design and/or delivery of the program activities. Finally, the program theory and logic models can be used to identify high priority research opportunities, such as the effectiveness of key program elements or additional market research, which can further inform the design and delivery of the program activities.[5]

[5] Additional information on program theory and logic models can be found in Rogers, Hacsi, Petrosino, and Huebner (2000), The TecMarket Works Team (2004), Frechtling (2007) and Knowlton, Wyatt and Phillips et al (2009).

Figure C-2 contains an illustrative program theory and logic model based on BGE's Large Commercial Prescriptive Program. This straw man example should be used to guide the structure and topics that should be covered in developing the program theory and logic models for other key programs selected by Navigant.

2.3.2 Evaluation of the Effectiveness of Campaigns to Increase Public Awareness of Energy Efficiency Opportunities

Studies by Itron to estimate the achievable savings potential suggest that the level of customer awareness of energy efficiency and program opportunities is often a key indicator of future program success.[6] The level of awareness of program opportunities in Maryland is likely to be low after ten years in which no significant programs were offered by most Maryland utilities. This places a premium on evaluating the effectiveness of the program and or social marketing strategies currently being employed in Maryland to increase customer awareness of energy efficiency and program opportunities.[7]

2.4 Provide Estimates of Historic Peak-load Impacts and Forecasts of Future Impact from Demand Response Programs

The Maryland utilities are investing substantial resources to solicit participation in demand response programs whose objective is to reduce peak demand on key high-load or high-cost days. Based on review of the demand response programs currently being offered in Maryland and the types of evaluation that have been conducted for these programs (see Appendix B), in this section we provide recommendations for evaluation planning and reporting on a going forward basis.

Maryland utilities should provide estimates of the level of demand response resources that were both available and called in the previous summer season in their annual program reports due to the Maryland PSC on January 30 of each year. In addition to documenting, or evaluating, the historical performance of demand response programs in the most recent year, the utilities should develop forecasts of expected load impacts in future years.[8] These forecasts should take into account information on load impacts at the per-participant level obtained in recent impact evaluations, and forecasts of expected enrollment in future years.

---

6 Itron has conducted numerous studies of energy savings potential, including most recently Florida (2009), Texas (2009), and the largest utility in Utah (2007).

7 A full scale, statewide market study of the awareness of energy efficiency options across all market actors is beyond the scope of the EmPOWER program evaluation effort. Programmatic marketing activities will be addressed by the evaluation team, but a systematic state-wide effort of public awareness is not contemplated.

8 In the California protocols, this type of analysis is referred to as *ex-ante* load impact evaluation.

Consideration should also be given to ranges of uncertainty around the forecasts due to uncertainty regarding per-participant load impacts, weather conditions and enrollments. These forecasts of the expected availability of demand response resources for the previous and following summer season (four months after February 1) should be presented in the annual program reports due January 30 of each year.

2.4.1 Utility Demand Response Evaluation Plans

Maryland utilities are expected to submit current or updated evaluation plans for their 2010 demand response programs in July. The utilities should include the following types of information in these evaluation plans:

- The methods that utilities propose to use to measure, evaluate, or verify the hourly load impacts (*e.g.*, load reductions in MWh/hour or MW) that were achieved in 2010 and 2011, for each event and each program. If they plan to apply per-participant load impacts from previous evaluations, they should address the issue of when future new evaluations will be conducted to update these estimates.
- The data that utilities plan to use to conduct those evaluations.
- The information that utilities plan to report (*e.g.*, program enrollment, estimated hourly load impacts for each event, average hourly load impacts per participant).
- Forecasts of enrollments and load impacts by program for the next three to five years, including assessment of key uncertainties in the forecast (*e.g.*, provide confidence intervals around the expected load impacts and document sources of uncertainty in enrollment and load impact forecasts).
- Discussion of plans for AMI installation, including dates of rollout, and plans for offering dynamic, time-based retail rates to various classes of customers (if the PSC approves any utilities' proposed AMI installations in the next three years).
- Discussion of the effect of plans for AMI and dynamic pricing on demand response evaluation plans.
- Discussion of how the evaluation plans meet the needs of the PSC to document progress toward EmPOWER Maryland goals and PJM requirements for demand response program evaluation and reporting.
- Discussion of the schedule for meeting the January 30 deadline for reporting annual program load impacts for the previous summer, and forecast load impacts for the following summer.

2.5 Cost Effectiveness of EmPOWER Program Portfolios

One of the primary objectives of the statewide EmPOWER program evaluation is to determine whether the portfolios of programs administered by MD utilities are cost effective. The PSC has not provided guidance on when utilities are expected to provide estimates of program or portfolio cost effectiveness as part of the statewide evaluation. Perhaps as a result, the Maryland utilities have not taken any steps yet to ensure the outputs of their program tracking systems can be readily input to their cost-effectiveness tools (see Appendix A for a summary and assessment of the utilities' current cost-effectiveness tools). Calculating the cost effectiveness of programs and portfolios will be more time consuming and difficult if it must be done manually.

Given the priority the Maryland General Assembly has placed on measuring the cost effectiveness of the EmPOWER programs, the independent evaluator originally proposed that the utilities provide cost-effectiveness estimates in their 2010 annual reports (due February 1, 2011). The utilities expressed concerns about the level of effort and cost that would be needed to complete this analysis and how to report results in way that is meaningful and actionable. In subsequent discussions between the independent evaluator, the utilities and PSC staff, we were able to develop a plan for estimating and reporting cost-effectiveness that seemed reasonable and could be supported by all of the participants.

This section provides background information and various options for completion of the cost-effectiveness analysis and concludes with the consensus recommendations based on discussions between the utilities, the independent evaluator and PSC staff.

2.5.1 Scope

The overall method of the statewide evaluation is to survey samples of program participants to develop realization rates (percent of utility-reported savings that can be verified through sampling). The realization rates for each utility program will then be applied to the reported savings from the entire program. To be able to explain and understand any realization rates that do not equal unity, the statewide evaluator will need to examine the algorithms and ex ante values used to generate the initial program savings forecasts in detail. Most of the algorithms and ex ante values being used by the utilities to report program savings are the same as those in the cost-effectiveness calculators that were used to develop program plans approved by PSC in 2009.[9] Thus it makes sense to review them both simultaneously. This review will also help understand any likely changes in the cost effectiveness of programs or portfolios.

[9] Except for AP, the utilities' CE tools were developed by ICF. Based on preliminary review the tools and assumptions that were used generally seemed reasonable.

The result of this review of ex ante values can be combined with the results from the statewide evaluator's estimates of program realization rates to produce a revised estimate of the cost effectiveness of the 2010 programs. However, making these changes to the inputs to the cost-effectiveness tools could require significant effort by the utilities and their contractors over the coming months.

There are at least two options for updating the cost-effectiveness tools to allow them to be used to estimate cost effectiveness based on the 2009-10 program savings:

1) Comprehensive Update - Update ex ante values in the cost-effectiveness calculators for which better information has been developed as a result of the utility statewide evaluator or independent evaluation effort as of December 1, 2010, or

2) Minimal Update - Use the old ex ante values from the 2008-09 cost effectiveness filings and simply change the participant and measure implementation assumptions to reflect 2010 reported data from the program tracking data bases.

The table below identifies the parameters that would need to be updated.

Table 2-1: Cost-Effectiveness Components that could be Updated for January 31, 2011 Filing

		Option A	Option B
		Minimal update for 2010 program based on program costs and participation levels only	**Comprehensive update using latest available data for 2010 program**
Benefits Parameters	Energy Savings per Unit	2008 Forecast	2010 TRM & Verified and or Using Realization Rate Approach
	Participants, Measures installed	2010 Reported	2010 Reported
	Avoided Energy Costs	2008 Forecast	2010 Forecast
	Avoided Peak Cost Forecast	2008 Forecast	2010 Forecast
	Environ Values	2008 Forecast	2010 Forecast
	Useful lives	2008 Forecast	2010 TRM
Cost Parameters	Program Costs-	2010 Reported	2010 Reported
	Rebates	2010 Reported	2010 Reported
	Incremental Measure Costs	2008 Forecast	2010 TRM & verified
Discount Rates	2008 Forecast	2008 Forecast	Revised to a common assumption, eg cost of capital, social discount rate
Key			
Forecast	Forecast values used in 2009-2011 plan		
Reported	Reported Values from 2010 tracking data bases		
TRM	Updated values using TRM review of MD TRM		
TRM & Verified	Updated values using TRM review and Evaluator findings as of 12/1/2010		

The relative cost of these two options is not entirely clear, but the utilities estimate costs of several tens of thousands of dollars to update the ex ante values in their cost-effectiveness tools.

2.5.2 Schedule

The burden on utilities, their contractors and Navigant will depend in part on the schedule for estimating cost effectiveness and scheduling of competing tasks. Developing realization rates based on participation through the third quarter (i.e., September 30, 2010) has been proposed to give more time to verify and true up evaluation results for use in the utilities' 2010 annual reports. This schedule would presumably free up some time for utilities, contractors and

Navigant to focus on developing robust cost-effectiveness estimates based on this initial 9 months of data.

Closing out the evaluation at the end of the third quarter would still allow estimates of savings for the full year (through December), but estimates for the last three months would be based on realization rates for program activity through September. In other words, the 9-month realization rates would be applied to 12 months of utility reported participation and measure installations.

Note: PSC staff may request preliminary estimates of cost effectiveness in early January 2011 based on the minimal-update option described in the reporting section. These cost effectiveness estimates would be based on participant counts and measure installations through September 30, 2010 and the ex ante savings values from the original 2008 cost-effectiveness tools.

At least three concerns arise around the prospect of using realization rates based on three quarters of program data with a September 30 close out date. First, given the slow ramp up of programs, estimated realization rates for the smaller utilities could be lower for the 9-month period than for a 10 or 11-month period. Consequently, the cost-effectiveness estimates could be lower.

Second, given the slow ramp-up of some programs, participation rates for those programs may be too low for meaningful sampling. Shortening the evaluation year by one or two months could further reduce the participation and thus sampling prospects. Consequently, there could be no verified evaluation or cost effectiveness results for some programs.

Third, the statewide evaluation plans will not be completed until July or maybe August, thus a three-quarter evaluation or close out would allow at most two months, and probably much less, for the statewide evaluator to sample and survey participants.

2.5.3 Reporting

Reporting cost-effectiveness estimates for some programs in early 2011 could be misleading or premature, given that most of the EmPOWER Maryland programs are new and program start-up costs will be spread across relatively small number of installations. The challenge is to provide meaningful cost-effectiveness estimates that can be used to guide policymakers and program implementers. We would not base the cost effectiveness of a power plant based on one year of operation; we should not evaluate the cost effectiveness of the EmPOWER Maryland programs based on one year of operation either.

Analogous to a power plant, the fixed start-up costs associated with the programs could be spread across the annual savings projected over the expected lifetime of the programs.

Unfortunately, programs are easier to mothball than power plants thus the lifetime of the programs is more difficult to predict. The EmPOWER Maryland programs are approved in three-year cycles, however, so one way to address the uncertainty about program duration and resulting benefits would be to estimate and report cost effectiveness for three scenarios:

1) Programs results for 2009 and 2010 only,
2) Programs continue through 2011 (three years of programs), and
3) Programs continue through 2015 (six years of programs).

Developing these scenarios would require applying the verified estimates of savings per participant or measure to forecasted participation and measure installations. Careful development and review of the forecasts of future program savings and costs after 2010 would be critical for developing meaningful cost effectiveness estimates. Hence, reporting estimates for the three scenarios could require somewhat greater effort and cost.

Reporting cost effectiveness this way would, however, underscore to stakeholders that the cost effectiveness of programs hinges in part on the duration of the programs and would provide more meaningful cost-effectiveness estimates that parallel those by policymakers to decide supply side investments.

2.5.4 Recommendations

Maryland PSC Staff and the independent evaluator have worked with the utilities and other stakeholders to develop a practical schedule and meaningful scope for the cost-effectiveness estimates. Specifically, we recommend the following:

At least two cost-effectiveness estimates should be provided: 1) estimates based on 2009-2010 actual participation rates for each program, and 2) estimates for 2009-2011 programs, which will be based on 2009-10 actual participation rates and 2011 projected participation rates.

Utilities should develop updated forecasts of program participation for 2010-2011 to be used in this cost-effectiveness analysis. The independent and statewide evaluators will review these projections to assess whether they are reasonable and sufficiently documented.

During the course of their desk review of savings estimates, the statewide evaluator should review and recommend revisions to the ex ante savings assumptions and algorithms in the existing cost-effectiveness tools and/or tracking systems based on their review of the Maryland TRM, findings from the Maryland baseline study, and other relevant studies. These recommendations should be shared and discussed with the utilities and the independent evaluator.

The independent evaluator should review both the findings from the statewide evaluator on measure level savings estimates and the program-level methodologies used by each utility to develop parameters for the cost-effectiveness analysis, including: incentive and administrative cost allocation, bundling of measures into programs, cost allocation for measures that save electricity or natural gas, company and customer discount rates, and common assumptions. Based on this review, the independent evaluator should recommend changes to the savings assumptions or methodologies that will be used to calculate cost-effectiveness estimates in its first verification report due on January 10, 2011.

The utilities, with cooperation from their own internal contractors and/or the statewide evaluator, should provide all data necessary to assess cost effectiveness, including collecting data participant measure costs, program costs, and remaining useful lives of measures.

Program and portfolio level cost effectiveness should be reported for each utility service area. The independent evaluator will be responsible for rolling up the results from all the utilities into a statewide total if this is requested by the Maryland PSC.

Realization rates used in the cost-effectiveness analysis should be developed based on program activity through September 30, 2010 and used in the initial estimates of portfolio cost effectiveness for 2009 and 2010 programs. These same realization rates should be applied to actual program activity through December 31, 2010 to produce annual estimates of program savings for 2010.

The utilities should provide cost-effectiveness estimates for their major programs and portfolios by March 15, 2011. These estimates would be based on the full panoply of findings from the statewide and independent evaluators for the 2009-10 program activities and include fully updated ex ante values.

Alternatively, if the Commission deems it necessary, the utilities could provide preliminary cost-effectiveness estimates February 1, 2010 as part of their annual reports and then final estimates May 15, 2011 in time for their 2012-2014 program plans. The preliminary February 1 cost-effectiveness estimates would include participation and installation counts through December 31, 2010, but would be based on the ex ante assumptions that were used in the utilities' 2009-12 program plans. The final estimates on May 15 would incorporate the full panoply of findings from the statewide and independent evaluators for the 2009-10 program activities and include fully updated ex ante values. They could also include updated measure costs, expected lifetimes, avoided costs and other assumptions from the Maryland Baseline Study and other sources.

2.6 Develop a Strategy to Begin to Measure the Net Energy Savings Attributable to Energy Efficiency Programs

Estimating net savings from programs presents significant challenges for evaluators and program administrators. The analytical challenges have been exacerbated in recent years as the number and size of program and policy activities have increased along with the number and variety of agents involved in program and policy delivery. Further confounding the estimation of net effects is the growing use of market transformation programs, whose impacts on the availability and price of efficient products occur over a longer time horizon and therefore significantly impact estimates of net savings over time.

Nevertheless, attribution of savings to programs is a critical evaluation element, whether the objective of the evaluation is to estimate program or portfolio cost effectiveness, provide inputs to utility system planner, determine compliance with EmPOWER Maryland commitments, or provide formative guidance to program implementers. While the levels of certainty and precision required for these various objectives may differ, in all cases, ignoring free riders and spillover can lead to perverse policy and program responses.

2.6.1 Recommendations

Net savings estimates are necessary to support the policy objectives related to determining compliance with the EmPOWER Maryland goals, determining cost-effectiveness at the portfolio and program level, and supporting improved program design. Specifically:

- Net savings should be an integral part of the 2009 and 2010 program evaluations. Using survey and other methods, the statewide evaluator should estimate free ridership and spillover associated with all of the sampled projects to develop a net-to-gross ratio for each evaluated program. Recent literature suggests that surveys with a 5-question battery may give the same results as more extensive surveys. The evaluation team is in the process of assessing these and other methods in order to reduce evaluation costs and survey fatigue.
- As with other evaluation elements, accuracy should come before precision. In other words, every effort should be taken to ensure minimal survey and other types of bias are present in the proposed evaluation method before increasing sample sizes in attempt to increase precision. The statistical precision levels associated with net savings estimates will necessarily be lower than gross savings estimates because the additional uncertainty associated with net methods must be combined with the uncertainty associated with the gross savings estimates. We recommend a minimum 90-30 (one-tailed) sample confidence and precisions level targets for combined net and gross evaluations.

- While survey devices will be the dominant approach used to estimate net savings, the statewide evaluator should employ a variety of approaches and survey a variety of participants (consumers, retailers, manufacturers) to gain multiple perspectives on what the net savings might be for high priority programs. We recommend the use of multiple approaches for high priority programs because this approach will allow evaluators the ability to triangulate net savings results for programs that represent significant fractions of the total portfolio savings, such as residential lighting programs.
- The statewide evaluator should implement a 2-stage evaluation approach for large I&C custom programs to minimize free ridership and support more accurate estimates of baseline energy usage in the absence of a program. The first stage of the evaluation should consist of a battery of questions designed to determine the natural or baseline rate of equipment replacement at the firm in question and the intentions of plant managers with respect to seeking out more efficient equipment before they were aware of the program offer. Our experience suggests the early interviews with key plant managers will significantly reduce the potential for an evaluation finding of a very low net to gross rate relative to what was anticipated by program designers.

2.7 Future Evaluation Challenges

We have deliberately steered away from recommending actions on some of the more controversial measurement problems in this section because we believe it will take time to build a robust evaluation structure that provides useful services to the program design and implementation communities in Maryland. However it would be a dereliction of duty not to at least foreshadow what some of these issues are likely to be and suggest they be reviewed again in mid 2012 after the first round of evaluations have established a track record. These issues are:

- Estimating the fuel substitution impacts of certain types of energy efficiency programs that promote both energy efficiency and other attributes that are consistent with the profit motive of the utility promoting them. Accounting for cross fuel impacts is tricky and complicated by any attempt to value the reduction of one fuel at the cost of increasing use of the alternative fuel.
- Estimating the remaining useful life of in situ industrial equipment within custom industrial programs and the standard efficiency levels of equipment that might have been purchased by industrial plant managers in the absence of a custom program. Advocates of this approach challenge the assumption that lifecycle savings for this program should be considered equal to a generic estimate of the useful life of the new replacement equipment multiplied by the annual savings. In many cases, the true lifecycle savings may only be equivalent to the remaining useful life of the equipment

times the first year saving, not the full useful life multiplied by the first year savings. The methods used to tease out estimates of the remaining useful life of industrial equipment from industrial plant managers are complicated and not yet universally accepted. Using this evaluation approach often has the effect of reducing the lifecycle savings in industrial applications by a factor of two or more.

- Estimating the effects of social marketing programs on customer behavior and or customer purchasing patterns.
- Estimating the effect of specific programs on "transforming" markets through the use of program success metrics or indicators.

We suspect that following through on the recommendations in Section 2.2 through 2.6 will provide more than enough challenges in the first two years and thus these issues can be deferred for now.

3

Guidance and Preliminary Plans for Deciding What Energy Efficiency Program Activities Should be Evaluated and the Effort to be Expended at the State, Utility and Program Levels

This section discusses the various tradeoffs that evaluators must consider when trying to increase the accuracy and precision of load impact estimates for energy efficiency programs at the statewide, utility area, and program levels. Based on its understanding of the Maryland PSC's priorities and those of the Empower Maryland utilities, the utility statewide evaluator, recommends various levels of precision and accuracy be applied at the portfolio, program and utility service area levels in the near term and over the course of the 2009 to 2011 program cycle. The independent evaluator will review these proposed confidence and precision levels within the context of the specific evaluation plans to be produced by the utility statewide evaluator in July 2010. In most cases, the proposed confidence levels from the utility statewide evaluator seem reasonable but will need to be re-evaluated and may need to be revised based on the relative priority of specific programs and the proposed emphasis on process compared to impact evaluations in the final evaluation plans.

3.1 Background

Any evaluation of energy efficiency programs will need to address challenges that may affect the accuracy and precision of the findings. This section provides a discussion of some of these challenges and the importance of transparently reporting evaluation methods and results to allow for these to be appropriately assessed by decision makers.

Many evaluations rely on surveys of customers and/or trade allies, as well as economic/engineering models, to produce estimates of program impacts. As a result, accuracy in the survey results and in the application of appropriate analytic methods is important. One way to view these challenges is to think about challenges to accuracy in the survey and analytic efforts separately. Table 3-1 illustrates one way to think about these different accuracy issues.

Table 3-1: Issues in Evaluation Accuracy

Issues in Survey and Data Collection Accuracy		Issues in Analytic Methods
Sampling: • Sample size • Resulting standard errors of the estimates	Non-Sampling: • Measurement errors • Non-response bias in survey implementation • Inappropriate (e.g., non-random) sampling methods. • Inappropriate determination of the sampling frame (e.g., faulty information on the population)	• Mis-specification of the model or analysis. • Violations of model assumptions (e.g., regression model assumptions, or such things as interactive effects in engineering models) • In-appropriate external validity assumptions (i.e., extrapolating model findings to an inappropriate set of customers, measures, or program metrics) • Modeler Error • Use of deterministic parameters and measurements when there may be uncertainty in these model constructs.

Note: This is not an exhaustive list, but factors like self-selection bias would be a violation of regression model assumption. This would also impact the external validity of the model and the extrapolation of model results to a larger population. Table 3-1 presents general categories of issues in assessing accuracy that embody a wide range of model specific issues.

In general, statistical precision is associated with the representativeness of the drawn sample while bias (the lack of accuracy) is associated with the other types of errors (bias in sample design, measurement errors, self-selection, and modeler error). Both statistical precision and bias are important to address in the design of an evaluation and in the presentation of evaluation results. Balancing the need for precision and accuracy subject to budget and data constraints is a challenge faced by every evaluation.

3.2 Multiple Objectives of the EE Evaluation Effort

This evaluation effort needs to meet multiple objectives held by different stakeholders, which creates additional evaluation challenges. This section discusses approaches and considerations important to balancing these evaluation objectives.

3.2.1 Tradeoffs in Meeting Statewide, Utility and Program-Specific Objectives

Energy efficiency program accomplishments from each administrator need to be estimated at the statewide level to determine if statewide energy savings goals have been met. In addition, ideally, load impacts for each program need to be measured to provide feedback to each program administrators. However, the limited budget will probably not allow the evaluation teams to meet all of these impact evaluation objectives and the process evaluation objectives discussed earlier. This section is about issues and options that need to that need to be considered when making these tradeoffs.

There will be economies in conducting a statewide evaluation of a portfolio of programs designed to meet statewide targets. However, disparities in the size of the Maryland utilities, as well as disparities in the contributions expected from specific programs will also create

challenges. Producing savings estimates at each of these levels of disaggregation will result in different requisite levels of confidence and precision. Table 3-2 illustrates some of these differences by displaying the distribution of expected (utility forecasted) energy savings across program administrators and program types across the three years energy saving targets. This table is based on information filed by the Maryland utilities with PJM in April of 2010 and is illustrative only. These savings estimates and the programs included in Table 3-2 are slightly different than the total program energy savings numbers filed with the Maryland PSC. As such, Table 3-2 may produce slightly different fractions of total savings by utility program and type.

Table 3-2: Distribution of Targeted MWh by Utility and Program Type

Target MWh %	Allegheny	BGE	DPL	PEPCO	SMECO	Total
Res. Audits	1.0%	0.0%	0.9%	4.6%	0.1%	**6.6%**
Res. HVAC	0.1%	1.0%	0.4%	2.0%	0.3%	**3.7%**
Res. Ltg/Appliances	1.2%	19.1%	1.6%	8.4%	1.5%	**31.8%**
Res. Low Income	0.1%	0.3%	0.2%	0.6%	0.1%	**1.3%**
Res. New Constr.	0.0%	0.3%	0.0%	0.0%	0.0%	**0.4%**
C&I Prescriptive	1.5%	14.1%	1.7%	10.4%	0.9%	**28.7%**
C&I Custom	0.4%	13.8%	1.2%	6.3%	0.9%	**22.6%**
C&I Small Business	0.0%	4.9%	0.0%	0.0%	0.0%	**4.9%**
Total	**4.4%**	**53.6%**	**5.9%**	**32.4%**	**3.8%**	**100.0%**

Source: April 2010 Updated Utility Evaluation Plans Prepared for PJM.

Table 3-2 also shows a wide variation in expected savings across program types, with three program types accounting for over 83% of statewide MWh savings. Each of these programs individually accounts for over 20% of the statewide savings while all but one of the other programs account for less than five percent of savings.

The skewed and somewhat disproportionate distribution of utility forecasted savings for specific program types and administrators illustrate the challenges of planning an evaluation effort that spans five different utilities rather than developing plans for a single entity in isolation with a given budget. One of the key challenges relates to the fact that the statewide evaluation team needs to decide how much of the total evaluation effort or dollars should be devoted to estimating the load impacts of SMECO programs when in total they are expected to count for less than four percent of the state-wide goals. Is it more important to have accurate estimates of the savings represented by the sum total of all utility programs or an accurate estimate of the savings for each utility area? It is expected that each utility will want to have an estimate of its program impacts and contribution to meeting the statewide goals, but there is a need to balance out the efforts to meet multiple objectives. To help make these tradeoffs, it is important to identify evaluation costs at the utility and program level so that

the relative evaluation expenditures can be estimated and compared (e.g., in terms of evaluation dollars per claimed MWh saved or an expenditure metric normalized to the overall evaluation cost). In this way, one can readily see whether proposed expenditures for particular program are many times more or less than that programs relative contribution to total savings at either the utility or statewide level.

3.2.2 Tradeoffs in Meeting State-wide and Utility-wide MWh Estimates with PJM Estimates for MW Reductions during a Four-Hour Performance Period

Another important evaluation objective is to support the utilities' bids of permanent capacity reductions into PJM's capacity market. The PJM is interested only in hourly reductions in energy use between 2 p.m. and 6 p.m. from June 1 through August 31. PJM requires that for all bids there be a 90% confidence that the MW reductions exceed a limit set 10% below the nominated value.

Developing estimates of reductions for given, short intervals of time pose different evaluation challenges and require different methods than procedures used to estimate monthly or annual MWh reductions. Allocating savings into tight time frames increases the need for precision in estimation. As a result, daily and hourly load shape analyses will need to be conducted to meet PJM requirements. This likely would not otherwise be necessary for all programs in the EmPOWER Maryland portfolio.

The need to complete two different sets of evaluations – one to meet the PJM peak performance hours' estimates and a second to check compliance with the EmPOWER Maryland energy efficiency goals – will require decisions regarding allocation of evaluation resources. There will be an overlap in the evaluation methods used, but isolating a four-hour contribution to MW reduction over a three-month period poses a different set of evaluation issues requiring additional analytics.

Discussions are ongoing between the statewide evaluator, utilities, and the PJM regarding strategies for nominating energy efficiency resources into PJM forward capacity markets. PJM states that the required level of confidence and precision for MW reductions within this four hour performance period must be at 90% confidence and 10% precision levels based on a one-tailed hypothesis test (i.e., there is to be a 90% confidence that the MW reductions exceed a limit set 10% below the nominated value). Obtaining this confidence and precision level will be challenging given that all the savings have to be precisely estimated within the targeted four-hour period. The initial evaluation work supporting nominated energy efficiency resource values and PJM's approach to assessing evaluation work will determine the level of effort. As yet, there have been no deliveries of savings into PJM's capacity market (but there have been preliminary nominations). The first opportunity to deliver

capacity is the summer of 2011. Until capacity nominations are made and PJM accepts the supporting evaluations, there will be some uncertainty about the requirements of this process.

The bidding of energy efficiency resources is new for PJM and new for the utilities, and there are still questions about how the process will work as well as the ability of the utilities to complete the PJM-required "post-installation M&V reporting" by May 1 for any EE resources bid into the capacity market for the upcoming summer. In addition, PJM retains the right to conduct, at the provider's expense, its own audits of the EE resources before or during the performance period:

> *"PJM or an independent third-party (as directed by PJM) may conduct an audit, at the EE Resource Provider's expense, of the energy efficiency installation prior to or during the Delivery Year. The M&V Audit may be conducted any time, including during the defined EE Performance Hours."*

As this process is new for all parties, there is not complete clarity over what information might be required of Maryland utilities to support a potentially expensive PJM audit of a utility's EE resource bids. Regardless of the 90% confidence and 10% precision requirements, the PJM might, due to concerns about potential bias, accept a lower, more conservative level of net MW during the 4-hour performance period.

3.3 Guidance on Program Evaluation Activities

In this section, we propose specific guidance related to the precision and accuracy targets to be used by the statewide evaluator at the statewide and utility program level. It is important to emphasize that the guidance developed here is based on the most current information from the tracking databases and new information on actual program participants may require changes in real time to the evaluation plan. There are many challenges associated with the development of energy efficiency programs, including program design, incentive levels, marketing strategies, participant enrollment, fulfillment (delivering the product or service), and various financial and administrative functions. Some programs in early years may perform better than expected, while some may not fully meet expectations. In addition, program tracking systems that record all of the baseline and installation data needed by evaluators are only now being developed by the Maryland program administrators. As a result, we anticipate it will be necessary to review the guidance provided here on a periodic basis as utility programs continue to be rolled out and ramped up during the second half of 2010. The purpose of this section is not to be prescriptive, but rather to set out the initial guidance to the evaluation team on the minimum precision targets that should be met or exceeded where possible and practical at the program and portfolio level.

3.3.1 Guiding Principles to be used in Determining Levels of Precision

Based on discussions between the independent evaluator and statewide evaluator, the statewide evaluator has agreed to abide by the following evaluation principles:

- The statewide evaluator's overarching goals will be to measure and verify the EmPOWER Maryland programs' contributions to the MWh and MW targets, provide accurate information to support good decision making with respect to investments in energy efficiency programs and resources, and support the utilities' bids of energy efficiency program capacity resources into the PJM forward capacity markets. The methods used by the statewide evaluator will generally focus on leveraging information in the program tracking systems to produce estimates of program impacts using a realization rate approach; i.e., an approach that assesses the fraction of the energy savings recorded in the tracking system that can be verified by verification/evaluation methods. In addition, this approach will ideally focus on providing feedback on specific aspects of the ex ante estimates that need to be adjusted (e.g., hours of operation, power reduction estimates, baseline parameters, net-to-gross ratios, etc.) to improve the accuracy of reported savings and future realization rates.
- The evaluation approaches used by the statewide evaluator will incorporate available information from the evaluation literature and work performed in other jurisdictions deemed applicable to the Maryland EmPOWER programs (e.g., the MD Baseline study and NEEP TRM).
- A transition to measuring net program savings in addition to gross savings is needed. The statewide evaluator is developing net-to-gross (NTG) methods based on a review of current methods used in Wisconsin, Massachusetts, California, and the Pacific Northwest.[1] It is likely that different net-to-gross methods will be used for different programs depending on the market actors involved; i.e., a one-size fits all approach (which has been used in some other jurisdictions) will likely not be employed.
- Confidence and precision target levels shall be as follows:
  - The statewide evaluator will target statewide portfolio level program impact estimates at 90% confidence and 10% precision levels based on a one-tailed test, for MWh targets, and use reference load shape analysis to estimate MW reductions in peak demand. This level of confidence and precision is the same as PJM's for energy efficiency resources delivered during the performance hours. Having estimates that exceed a limit that is 10 percent lower than the target with

[1] These methods will need to be compared to the methods recommended in the soon to be released report on net to gross methods commissioned by the North East Energy Efficiency Partnership a few months ago.

a probability of 90 percent would seem to provide appropriate information for policy makers. An 80/10% two-tailed test would, for example, indicate that an energy efficiency program expected to achieve a savings of 1,000 kWh would have to have an estimated value that falls between 900 kWh and 1,100 kWh. A one-tailed test simply states that there is a 90% probability that impacts exceed 900kWh. Given uncertainties in resource planning across all utility resources (e.g., costs of supply-side and renewable resources); this is proposed as a reasonable level of statistical confidence and precision for the statewide goals.

- For individual program categories at the statewide level, the target will be 90/20, based on a one-tailed test, for high priority programs (high priority to be defined based on contribution to reaching targets or potential for future contributions). Minimum 90/20 confidence/precision levels would be used to address programs that in aggregate account for approximately 80% of the total portfolio impacts.
- For individual programs in the statewide portfolio that contribute smaller amounts to the savings targets, the impact estimates will be assessed with an 80/20 one-tailed test. These programs will also be examined for their ability to make larger contributions in the future through formative process evaluations.
- The focus of the statewide evaluator's efforts will be on estimating program impacts in 2010 and 2011 as these impacts will provide the greatest contributions to the EmPOWER Maryland targets. Impacts from 2009 program efforts will be estimated by drawing samples and using secondary research to test whether the 2009 impacts recorded in the tracking system are found to be statistically different from the 2010 and 2011 impacts. This statistical approach will be used to develop adjustment factors for 2009 program impacts. This approach will focus the evaluation resources on the most current and important research questions.
- Research priorities will need to be developed. In 2010, the statewide evaluator will attempt to address programs contributing 80% of the impacts in the tracking systems. The statewide evaluator's efforts will include participant research, deemed savings (from secondary sources), and research from other evaluations judged to be applicable to the MD EmPOWER programs.
- All programs at all utilities will, at a minimum, undergo a desk engineering review, including review of tracking system impact calculators and secondary data, to develop realization rates and program impacts for every program for every utility. The Mid-Atlantic Technical Resource Manual (TRM) along with neighboring state TRMs (e.g., Ohio and Pennsylvania) will be used as source documents for these estimates.

Given the early stage of many utility EE programs, the statewide evaluator will work with utilities to determine how to prioritize and allocate evaluation resources between process and

load impact evaluations for each of the major programs. These program specific allocations will be provided in the July evaluation plans based on program information needs judged as important by the utility and the independent evaluator.

The independent evaluator has discussed these proposed confidence and precision levels with the utility statewide evaluator and agrees that they represent a useful set of minimum precision targets to use in scoping out the draft evaluation plan for Maryland in July. However, the independent evaluator and PSC reserve the right to request that precision levels are either increased or decreased after the data on the latest program savings share of total savings is available and the major tradeoffs between allocating resources between process and load impact evaluations have been made by the utility statewide evaluator within the available budget.

3.3.2 Additional Discussion and Guidance Related to Precision and Accuracy

Additional guidance on how the statewide evaluation team plans to develop evaluation plans at the statewide and utility service area levels is provided below. In this section, we deal with the other important factors such as the relative contribution of programs to the total portfolio savings that will shape the evaluation plans besides the sampling criteria and confidence levels discussed in the earlier sections. .

Program's Contribution to Portfolio Annual Impacts at the State-Wide Level

Approximately 80% of the energy savings and peak demand reductions should be evaluated using some form of in-field measurement for selected parameters, even if the impacts were initially estimated using a deemed savings approach (i.e., validating deemed estimates). Some of this in-field data is expected to come from the Baseline study. Data collected as part of the Baseline Study is needed for estimating baseline conditions for impact algorithms. Approximately 20% of the energy savings and peak demand reductions at the statewide portfolio level can be based on deemed savings or engineering algorithms. However the statewide team expects to perform a vigorous desk top review of the savings assumptions used in these lower priority programs and compare them against relevant secondary information.

Work on program evaluation should take into account the stage of program implementation and the overall expected impact of the program. If a program (e.g., residential lighting) is expected to produce a large fraction of the portfolio savings, it would be better to evaluate that program sooner rather than later to determine whether or not that program is on track to deliver the anticipated impacts. Similarly, it may be useful to examine important programs, in terms of expected overall impacts, even if the programs are at an early stage of roll out to ensure they reach their projections and make their contribution to the portfolio targets.

Table 3-3 and Table 3-4 show the expected contributions to the three-year portfolio MWh and MW targets by program across the five utilities.[2] The highlighted programs account for more than 80% of the contribution to portfolio targets. The tables also show that BGE and PEPCO programs account for the vast majority of contribution toward the targets – 86% of energy savings and 83% of peak demand reduction.

Table 3-3: Distribution of Targeted MWh by Utility and Program Type – High Priority MWh Programs Highlighted (80% Contribution to Statewide Target)

Target MWh %	Allegheny	BGE	DPL	PEPCO	SMECO	Total
Res. Audits	1.0%	0.0%	0.9%	4.6%	0.1%	**6.6%**
Res. HVAC	0.1%	1.0%	0.4%	2.0%	0.3%	**3.7%**
Res. Ltg/Appliances	1.2%	19.1%	1.6%	8.4%	1.5%	**31.8%**
Res. Low Income	0.1%	0.3%	0.2%	0.6%	0.1%	**1.3%**
Res. New Constr.	0.0%	0.3%	0.0%	0.0%	0.0%	**0.4%**
C&I Prescriptive	1.5%	14.1%	1.7%	10.4%	0.9%	**28.7%**
C&I Custom	0.4%	13.8%	1.2%	6.3%	0.9%	**22.6%**
C&I Small Business	0.0%	4.9%	0.0%	0.0%	0.0%	**4.9%**
Total	**4.4%**	**53.6%**	**5.9%**	**32.4%**	**3.8%**	**100.0%**

The three highlighted rows account for 83.1% of MWh savings

Source: April 2010 Updated Utility Evaluation Plans Prepared for PJM

[2] Note Table 3-3 is a copy of the earlier Table 3-2 found earlier in this section but contains additional highlighting and details.

Table 3-4: Distribution of Targeted MWh by Utility and Program Type. High Priority MW Programs Highlighted (80% Contribution to Statewide Target)

Target MWh %	Allegheny	BGE	DPL	PEPCO	SMECO	Total
Res. Audits	1.6%	0.0%	1.1%	5.5%	0.2%	**8.4%**
Res. HVAC	0.3%	1.8%	0.9%	4.7%	0.6%	**8.3%**
Res. Ltg/Appliances	2.9%	20.6%	1.6%	8.7%	1.0%	**34.8%**
Res. Low Income	0.1%	2.2%	0.2%	0.5%	0.1%	**3.1%**
Res. New Constr.	0.0%	1.0%	0.0%	0.0%	0.0%	**1.1%**
C&I Prescriptive	2.2%	10.9%	1.5%	8.9%	0.8%	**24.3%**
C&I Custom	0.5%	5.7%	1.1%	5.5%	0.7%	**13.5%**
C&I Small Business	0.0%	6.5%	0.0%	0.0%	0.0%	**6.5%**
Total	**7.6%**	**48.8%**	**6.3%**	**33.9%**	**3.5%**	**100.0%**

The four highlighted rows account for 81% of MW reductions.

NOTE: HVAC is proposed by the utility statewide evaluator to be selected over residential audits due to a greater ability to verify savings and develop realization rates at a lower cost than residential audits. This will be assessed further during review of the draft utility evaluation plans.

Source: April 2010 Updated Utility Evaluation Plans Prepared for PJM.

It will be important to clarify with the utility statewide evaluators which of the three different types of EmPOWER Maryland audits are currently included in the "Residential Audits" category in this table. Residential audits have a low share of total program savings here (8.4%) but a much higher share of total reported program participation and savings in the first quarter of 2010. The distinction between savings produced by audit recommendations and the savings associated with low-cost and no-cost measures installed during the audit may need further clarification.

In-field data collection should begin in the summer of 2010 for the major programs if at all possible. If the program sampling frames are not available in time for summer data-collection data from 2010 in-field work, the limited work performed this summer may need to be combined statistically with samples from 2009 and 2011 to revise the three year "cycle realization rates."

Need to Assess Cost-Effectiveness at the Program and Portfolio Level

Cost-effectiveness analysis is not a component of the utility statewide evaluator's scope of work; however, it is recognized that realization rates and net savings estimates will be components of any cost-effectiveness assessment. The statewide evaluator should provide all data necessary to assess cost-effectiveness, including any participant cost data it collects during the course of its evaluation work. Generally, the task of verifying participant cost data has not been a component of traditional impact or process evaluation, however if utility statewide evaluator does not perform this task, the utility program administrators will still be

responsible for either collecting or estimating this data, since it is necessary input to total resource and societal cost tests. The utility administrators, who will be responsible for collecting this incremental cost data, should inform the Maryland PSC if it is determined that the utility statewide evaluator will not be involved in this data collection effort.

Data from the evaluation will be provided to support cost-effectiveness analysis at the program level both statewide and by utility, although the utilities with the smaller programs will likely not have data collected at the same level of confidence and precision as will be the case for BGE and PEPCO.

Verification Activities That May Be Needed to Refine Deemed Savings Estimates and Savings from Custom Measures

Given that an initial Maryland Technical Reference Manual (TRM) (produced by the Northeast Energy Efficiency Partnership - NEEP) was just recently completed and only covers a fraction of the measures offered in Maryland utility programs, it is not yet clear the role that deemed savings will play in the statewide evaluation. It is expected that deemed savings will be available for use by evaluators for some measures using the TRM developed by NEEP for the Mid-Atlantic region and that this data may be supplemented by other sources (such as the TRM research in Pennsylvania and other jurisdictions). Over the course of the evaluation, the statewide evaluator will check deemed savings estimates against both baseline data and selected in-field measurements to validate the accuracy of deemed savings estimates for substantive measures. Measures that account for less than three percent of a program's impacts may not warrant substantive field research; thus, secondary research may suffice for validating some deemed savings estimates.

Type and Depth of Evaluation Appropriate to Each Type of Program

As discussed above, the programs that account for 80% of energy impacts should be the initial focus of the evaluation. In addition, programs that are candidates for bidding peak demand reduction into the 2011 PJM market will be given priority for in-field summer data collection. There is a requirement that all impacts bid have in-field analyses conducted prior to May 1 of the corresponding delivery year. The PJM peak period covers summer months, and collecting data on programs believed to meet the PJM requirements may need to have some in-field data collection in August of 2010, if possible given time and data constraints.

In general, programs should undergo different levels of evaluation review. A nested approach will be used with all programs classified into a level of analysis:

- **Level 1** – Technical review of tracking system program and measure estimates. In some cases, this will involve the review of deemed savings that are embedded in the

tracking systems. This analysis is expected to be performed for all utilities and all utility programs.

- **Level 2** – In-field analysis of assumptions in tracking systems that may involve data collection with program participants and/or trade allies.
- **Level 3** – Calibrated simulation and engineering analyses using in-field data to develop participant and measure-specific analyses. Analysis of changes in consumption using econometric and billing analysis methods (may include participant and non-participant data and customer billing data from representative periods before and after installation of efficiency measures).

Some programs may have a Level 1 analysis in 2010 and a Level 2 analysis in 2011. Impact estimates for programs implemented in 2009 will be developed based on an analysis of 2010 participants and a targeted sample of participants/measures installed in 2009. This analysis will allow Navigant to determine if adjustments to the 2010 estimates are needed to develop 2009 program load impact estimates.

Staggered Evaluation Schedule to Stabilize Production and Review Processes

Evaluation efforts will need to be staggered to mirror the staggered launch and participation rates observed for the Maryland programs. A proposal for staggering research for specific programs will be provided by the statewide evaluator to ensure timely collection of information this summer for some key programs. The evaluators will focus their efforts on programs likely to yield impact estimates that are sufficiently reliable for use in bidding capacity into the PJM auction for the 2011 delivery year. The following considerations will be used in developing this staggered schedule:

- Different programs are in different stages of implementation, which will result in a rolling evaluation of programs addressing both impacts and processes.
- While the impact evaluation in particular is expected to be a rolling evaluation not tied to any specific 12-month period, reports will be prepared to comport with the required annual and semi-annual periods defined by the PSC and PJM.
- Using a rolling evaluation will allow impact evaluation to be better integrated with implementation - i.e., early identification of realization rates and net impacts to allow program implementers to use evaluation results to revise program designs as needed to achieve anticipated performance levels.
- The use of a rolling frame for evaluation will allow flexibility in the time periods used to report results so that different deliverables prepared by the statewide evaluator can meet reporting schedules with the best available information. Note: It is important to finalize a schedule of deliverables for the evaluation effort.

- There will be seasonal work elements specifically with respect to weather sensitive measures (e.g., HVAC and lighting and their impacts on peak demand) in each year with special emphasis on summer and winter impacts, as well as annual energy impacts.

4

Guidance on Contents of Utility Program Tracking Databases and Processes for Quality Control

The program tracking systems of the EmPOWER Maryland utilities are being developed and managed by a variety of contractors. BGE and SMECO have contracted with ICF to develop and maintain their respective program tracking systems and PHI (PEPCO and DPL) has contracted with a suite of contractors to do the same – namely Honeywell for residential programs, Lockheed Martin (LM) for industrial and commercial programs, and Comverge for demand response programs. During reviews of program tracking systems with Navigant and Itron, AP indicated interest in possibly outsourcing this work in the future and has since issued a request-for-proposals for this work.

Navigant and Itron initiated a series of meetings with each of the utilities to review the status and functionality of the utilities' program tracking systems. Initial meetings were held with each of the five utilities and their respective program implementation contractors. (A summary of the key points discussed at these meetings will be sent in a separate memo to the Maryland PSC and each of the Maryland utilities.) Navigant and Itron expect that an iterative process will be needed to fully understand the design and comprehensiveness of the utility tracking systems, one involving initial data transfers and then follow-up discussions to resolve questions and data gaps.

4.1 Inventory of Existing Utility Database Tracking Systems

At present, the utility program tracking systems exhibit a wide range of functionality and all appear to include (at a minimum) the ability to track program participants, record customer and trade ally contact information, collect information about incented efficiency measures, and provide basic reporting functions. Within utilities, variation exists in the level of functionality across programs – this seems to be largely a function of the age of the programs. Tracking systems for programs that have recently been launched tend to be less mature.

The extent to which the utilities' program tracking systems have actually been populated was not entirely clear during the meetings with the utilities and their implementation contractors. Navigant has made several data requests to the utilities for data to support the various

evaluation functions (i.e., impact, process, and market), and the extracts provided by the utilities will reveal the extent to which data fields have been populated. The extracts will also test the degree to which various query and reporting functions are operational.

The meetings revealed that energy and peak demand savings estimates are generally included in the utilities' program tracking systems.[3] In many instances, the assumptions and algorithms used to develop the estimates are contained in spreadsheets and other processing tools external to the tracking systems. Generally, the assumptions and algorithms are available electronically, but only Lockheed Martin provided tabs linking background documents with tracking system elements. Navigant has requested the assumptions and algorithms and supporting spreadsheets/processing tools from each of the utilities as these items will be central to many evaluation activities. Of critical importance will be assumptions made and field data collected by the utilities regarding the baseline conditions used to develop the energy savings estimates.

In all cases, the utilities and their implementation contractors indicated that savings calculators could be fairly easily updated to reflect changes in *ex ante* values or algorithms. Notably, all of the utilities indicated reluctance to change *ex ante* values in their calculators for the purposes of recalculating saving estimates from programs operated between 2009 and 2011 for the first EmPOWER program cycle. The utilities universally expect that the statewide evaluation results will be used to develop *ex ante* values for use in the 2012-2015 program plans, but that they will not be required to incorporate the new *ex ante* values into their 2009-2011 tracking systems in a retrospective manner.

The extent to which automatic quality control measures and procedures are incorporated into the utilities' program tracking systems appears mixed, with ICF claiming to have automatic validation and other quality control measures in place throughout its tracking systems and the other contractors stating that they primarily rely upon visual quality control and validation. Most (if not all) of the tracking systems contain unique identifiers for tracking different measures and customers and the frequency in which data is reportedly uploaded and updated (e.g., daily, monthly, etc.) seemed generally reasonable.

Finally, as noted previously, AP is the currently using internal staff and resources to develop a program tracking system. During the meeting with AP staff, it was revealed that AP is considering contracting out this task and AP has since issued an RFP for this work. AP requested guidance from Navigant and/or Itron on the design and functionality considerations that should be included in an RFP should the decision be made to outsource the tracking

[3] Honeywell was not aware until recently that it needed to be tracking kW impact as well as kWh – Navigant is working with Honeywell to develop calculators for kW impacts.

system. The tracking systems in place at the other utilities could be used as models for this type of information; however, the tracking systems are proprietary so caution will be needed in providing guidance to AP. Proprietary concerns could frustrate efforts over the longer term to create protocols for data collection and reporting information.

The utilities are currently in the process of collecting program tracking data through online and paper program applications. Much of the data being collecting by these program applications will need to be included in the program tracking systems to be used in the evaluation. At this time, the utilities are working to provide, but have not yet provided, actual program tracking system data/extracts in a usable format for evaluation. The utilities have provided demonstrations of the data available in the tracking systems and will be providing data dictionaries that describe the data being collected and their specifications within the tracking systems. Once Navigant and Itron have received the data dictionaries and reviewed the tracking system variables, the two evaluation teams will be able to develop a detailed list of the specific fields necessary for the evaluation. In addition, Navigant and Itron will be able to develop recommendations for additional fields to be added to the program tracking systems (if needed) that would assist in current and future evaluation efforts.

4.2 Assessment of Program Tracking System Capabilities

4.2.1 Data Being Collected and Matched to Evaluation Needs

Table 4-1 summarizes the results of our inventory of the data currently being collected in each of the utilities' program tracking systems.

Table 4-1: Review of Data Collection Capabilities of Each Tracking System

Data Collection Capabilities- Ability to Access:		BGE	Allegheny Power	PHI Residentia	PHI Com/Ind	SMECO
	Customer Contact information and record of customer participation flow- from application to eligibility decision to cutting check- for	Yes, but some information gathered on paper and	Partial, some info comes from third party vendors in	Yes, but 3 separate dB systems for EE Res, EE C/I and DR	Yes, but 3 separate dB systems for EE Res, EE C/I and DR	Not yet clear, should be same as BGE
	Billing and customer meter ID information from utility CIS system	Yes	NO, no links from spreadsheets to CIS	Yes	Yes	No
	Description of Equipment installed- capacity and eff iciency ratings being collected for:	All Programs	Some Programs	Some Programs	Some Programs	Some programs
	Description of Equipment replaced- capacity and efficiency ratings being collected for:	Most Program	Some Programs	Some Programs	Some Programs	
	Operating hours for participating customers?	Paper self report from customers, some defaults	Default value from PA TRM	Collected in paper customer apps in some cases, elec others	Collected in paper customer apps in some cases, elec others	Don't know yet
	Customer Reasons for non participation in program/recommended installs?	Not collected	Not collected	Not collected	Not collected	Not collected
	Installation/Purchase Cost of more efficient equipment?	All Programs?	Some Programs	Some Programs	Some Programs	Some Programs
	Building Type- ten bldg types for Com, three for residential- SF, MF, MH	Yes for C/I	Not being collected	Not being collected	Not being collected	Not being collected
	One integrated system or multiple DB for each sector or program?	One Integrated DB	Multiple DBs	Multiple DBs	Multiple DBs	One Integrated DB

Table 4-2 provides the same type of systematic assessment of the ability of each database or multiple databases to perform sorts or queries for analytical or planning purposes. The purpose of this table is to develop some perspective on what improvements might be needed to better support program planning, implementation and simply management efforts. For example, nearly all of the program administrators are striving to develop a program dashboard functionality that can quickly assess how reported savings compare to program savings goals on a weekly or even daily basis. Program administrators may also want to assess the feasibility of including a portfolio-level estimate of cost effectiveness as part of this dashboard.

Table 4-2: Query and Report Generation Capabilities of Program Tracking Systems

Query and Report Generation Capabilities-Ability to Sort Savings By:		BGE	Allegheny Power	PHI Residentia	PHI Com/Ind	SMECO
	-Sector?	Yes	Yes	Yes	Yes	Yes
	-End Use?	No but possib	No but possib	No but possib	No but possib	No but possible
	-Building Type?	Yes	No	No	No	No
	-Measure or Technology?	Yes	Under Construction-	Yes	Yes	Yes
	Parent Customer for chain accounts	Yes	Maybe	Yes	Yes	Not sure
	Can DB generate reports of cumulative customer and program expenditure by sector and county?	Feasible but not a preset report, need to add county codes	Probably Not until integrated system is set up	Feasible but not a preset report, need to add county codes	Feasible but not a preset report, need to add county codes	Feasible but not a preset report, need to
	Customized Report to populate quarterly templates from PSC?	yes in theory	Not at this time	Possible to program	Possible to program	Yes in theory
	Can Db Link participants to CIS system to estimate customer participation rates by program or sector?	yes in theory	Not at this time	Yes but would need more county code information from CIS system	Yes but would need more county code information from CIS system	Yes in theory
Tracking Progress-Dashboard available to compare reported savings to annual Empower savings goals?		Yes	Under Construction	Under Construction	Under Construction	Under Constructi on

Table 4-3 summarizes the ability of the tracking systems to support scenario analyses and to determine how sensitive estimates of reported savings are to changes in key assumptions.

Table 4-3: Scenario and Sensitivity Analysis Capabilities

What if Analysis: Does DB have capability to?		BGE	Allegheny Power	PHI Residential	PHI Com/Ind	SMECO
	Extrapolate current Participation trends to project year or cycle end savings	Not now but could program this	Not at this time, historic record is not sufficient	Possible to program	Possible to program	Not now but could program this
	Replace energy savings estimates with revised TRM estimates and determine delta savings	Yes but advise against early change out	Yes but may involve manual changes	Possible to program, might have to manually change for some algorithms	Possible to program, might have to manually change for some algorithms	Yes but advise against early change out
Quality Control						
	Manual or automatic QC-validation	Mostly manual	Manual	Mostly manual	Mostly manual	Not sure
	Range checks and alert indicators for all key cell entries?	Not currently	Not currently	Not currently	Not currently	Not currently
	Automatic Check for duplicate applications for same address	Not sure	Not currently	Not currently	Not currently	Not sure
	Independent entity to perform data entry review?	Yes	Not sure	Yes	Yes	Not sure
Possible functionalities to add to Data Base						
	Automatic Links to Cost Effectiveness Tools	No plans	No plans	No plans	No plans	No plans
	Projections of program savings or costs to GIS maps to show regional uptake rates	Not now but maybe later	No	No	Yes for C/I	Yes
	Analysis of effectiveness of Program Advertizing in Stimulating Applications	would occur outside of data base	would occur outside of data base	would occur outside of data base	would occur outside of data base	would occur outside of data base

4.3 Minimum Database Requirements to Support Evaluation

One of the purposes of a utility program tracking system is to provide quality data in a usable format to support program evaluation. To effectively serve this purpose, the program tracking system should be able to summarize the data for reporting purposes, extract the data necessary to conduct cost effectiveness analyses, and contain the data needed for use as sample frames for primary data collection efforts.

A single program tracking system should be provided by each utility containing all the necessary data (current and accurate) from each of their programs. The tracking system should be delivered as a relational database with primary keys to link tables as necessary.

Including all of the necessary data from each program under a standard field name greatly simplifies the production of summary statistics across programs.

To avoid duplicating the customer, contractor, payee, and project data associated with each record, a relational database structure is recommended in the following format:

- **Customer Data:** Unique ID to relate customer data to measure data, all associated account numbers for a given site, site contact data (name, address, phone), project contact data (name, address, phone),building type, sector, square feet, business hours
- **Contractor Data:** Unique ID to relate contractor data to a project, contact data (name, address, phone)
- **Measure Data:** Unique record ID, CustomerID linking to customer data, ProjectID linking to project data, PayeeID linking to Payee data, Program, delivery mechanism, measure name/description, measure code, end use, quantity, unit basis, *ex ante* estimates (with links to source of estimates), measure cost, baseline, hours of operation, effective useful life, *ex ante* net to gross ratio, incentive/unit
- **Payee data:** Unique payee ID, customer vs. contractor, paid date, incentive amount
- **Project Data:** Unique project ID, install date, status, ContractorID linking to contractor data
- **Program Administrators** should also consider coordinating the development of a common naming convention for measures to ensure the codes used in these program tracking databases are consistent with the measure codes being used in the baseline database construction effort being managed by BGE.

Additional fields that could be helpful for the evaluators may be suggested after the review of the data dictionaries. When a given field is not applicable to a particular measure it should be included and reported as "NA." This will allow all programs to exist in a single table. The source of *ex ante* savings estimates should also be provided with all relevant supporting documentation.

It is highly recommended that each record within the tracking system be assigned a unique identifier that remains with that record throughout the life of the program cycle. This will allow new records and changes in status from one release to the next to be easily identified.

The data in the program tracking systems should undergo quality control prior to being entered into the reporting systems. Quality control checks should be in place to ensure that all necessary fields are filled out in their entirety and to validate the data as being within an acceptable range. Some utilities have the capability built into their data entry processes that does not allow a user to submit an application or move to the next step without completing

all necessary fields with applicable and accurate data. This should be a goal for all utilities to work toward.

4.3.1 Suggested Long-Term Improvements in Tracking Systems

Additional data that utilities might consider tracking if they do not already do so includes complete baseline information for all applicable measures, hours of operation for all applicable measures, and business hours for each participating industrial and commercial site. A more complete list of additional data to track (if any) will be provided by Navigant and Itron after the review of the data dictionaries.

Suggested longer term improvements in the program tracking systems will likely include coordination across utilities to report common measure names and building types as well as consistent database tables and field names, to the extent possible. Greater consistency across utilities would enable common reports to be created and facilitate comparisons across utilities, programs, delivery mechanisms, end-uses, measures, and building types.

Another long-term goal for the utilities should be the incorporation of rigorous quality control processes in their respective program tracking systems. Whether program applicants are completing the forms online or utility/contractor staffs are manually entering data from paper applications, checks should exist to ensure complete and accurate data entry prior to accepting the application.

5

Guidance on Program Reporting

This section provides guidance on the use of standardized savings terms, the contents of the annual savings reports produced by the utility program administrators annually on January 30^{t} and the contents of the statewide evaluation report to be produced by Navigant for the first time in January 2011. We expect to provide our recommendations on the topics to be included in the Annual Verification Report (to be produced by Itron in cooperation with the Maryland PSC) after we have had a chance to determine what level of resources will be required to verify the savings appearing in this Annual Verification Report and to what extent our clients wish us to provide independent estimates of the employment or environmental impacts of the Maryland energy efficiency programs.

5.1 Standardize and Clarify Energy Savings Terms

Our review of program planning and reporting documents suggests there are some terms associated with different types and vintages of impact estimates that need to be standardized and clarified. Table 5-1 summarizes our recommended changes to each term and the rationale behind these recommendations.

Table 5-1: Standardized Savings Terms

Current Savings Terms	Revised Savings Terms	Rationale
Estimated	Forecast	"Estimated savings" which is used to describe ex ante savings estimated before the program launch, could be confused with ex post savings estimates conducted after programs are implemented.
Actual	Reported	Actual savings implies some level of verification; savings estimates are actually based on measure counts and deemed savings assumptions, not actual savings
No terms that match the concept of verified savings are found in Maryland reports	Verified	Third party evaluator verifies both participation in programs and energy savings per participant to derive ex post program savings estimates

In sum, we recommend that the utilities use these revised savings terms in their quarterly reports and their next annual saving report in January 2011.

5.2 Contents of the Annual (Reported) Savings Reports Produced by Program Administrators

The program administrators' annual savings reports should be the place where policy makers can find high-level summaries of impacts that were achieved in the previous year, and how the actual savings compare to previous year forecasts and overall EmPOWER Maryland savings goals. The reports should include a section stating whether the estimates of program savings made in the previous year have been verified by their independent evaluator.

We recommend the program administrators begin using reporting requirements tables recommended in the NEEP EMV Forum's draft report entitled *Common Statewide Energy Efficiency Reporting Guidelines* (May 2010).

In addition, Itron recommends that the reports include a discussion of:

- How coincident peak savings were derived,
- Uncertainty ranges associated with both individual program and portfolio savings estimates; and
- Proposed modifications supported by either evaluation studies or other factors or judgments from program administrators.

In addition, program administrators should provide their best estimates of the cost effectiveness of both their energy efficiency portfolio and their major programs.

Finally, as we recommended earlier, program administrators should provide estimates of the peak saving associated with their demand response programs in the previous summer season and a forecast of the level of average peak demand savings likely to be available in the upcoming summer season.

5.3 Contents of the Annual Verified Savings Report

We anticipate that Navigant will work with the Maryland utilities to design a comprehensive report summarizing the results of their load impact and process evaluations completed in 2010. Rather than providing a detailed outline of what that report should contain, Itron recommends the statewide evaluator and utilities make sure the report includes the following information:

- Realization rates (ratio of verified to reported savings) for each program and an explanation of why they are not equal to one.
- A summary of key program design changes recommended as a result of process evaluations or other market changes.
- A summary of the key load impacts studies completed and how they contributed to an estimate of energy and peak savings at the program and portfolio level.
- A discussion of how or whether the evaluations conducted can be used to confirm or verify the use of deemed savings estimates for some or all programs.

6

Proposed Planning, Evaluation, and Reporting Schedule for 2010-2012

This section develops a proposed schedule for the evaluation of energy efficiency and demand response programs operated by the EmPOWER Maryland utilities from 2009 to 2011. It includes a set of deliverables for program administrators, the statewide evaluator (managed by the Maryland utilities) and the independent evaluator (managed by the Maryland PSC), along with rationales for why these dates were chosen. This schedule, if approved, would replace/supplement the existing schedules for reporting program savings for the various entities involved with the review, approval and evaluation of the impacts of these programs

6.1 Scheduling Constraints and Reporting Deadlines which Shaped the Proposed Schedule

In order to develop a schedule that met the needs of a variety of organizations involved in this process, Itron consulted previous PSC orders and stakeholders in Maryland to identify the key milestones and deadlines that must be held constant or could not be changed, and considered when evaluation reports must be available to be the most useful to program managers and the Maryland PSC. We identified the following fixed schedule constraints that repeat on an annual basis:

- January 30 – Annual Programmatic Savings Report from Maryland utility administrators to the PSC.
- March 1 – Annual Report of the Maryland PSC to the Legislature on EmPOWER MD Programs
- May 15 – Post-installation EM&V report due to PJM.

Based on discussions with MD staff and stakeholders, we concluded that it would not be feasible to change any of these dates in the near term. As a result, one of our principal challenges was to develop a schedule that would provide an early readout on both reported and verified savings from the 2010 programs at a minimum and hopefully provide some level of verification for the 2009 programs at least one month before the legislative deadline of March 1 each year.

In addition to these fixed scheduling constraints, we considered the need to produce an annual evaluation of the energy and peak savings from these programs for each of the five Maryland utilities and to have Itron produce an independent verification of these savings estimates on an annual basis. Ideally, the savings results from both studies would be available a few weeks before the Maryland PSC is required to report on the savings and cost effectiveness of the EmPOWER Maryland programs.

Finally, we considered the need to make sure evaluation reports are both available and useful to the professionals working on the design of energy efficiency programs.

6.2 Options to Meet the Scheduling Constraints

The chief problem identified in developing a schedule to meet these needs was the ability of Itron and Navigant to produce reports of verified program savings within two months of the official closing date of the program year on December 31 of each year in time for a report to be sent to the PSC on or around February 1 to meet the March 1 legislative deadline. Elements of this problem include:

- It often takes utility program managers at least a month to clean and verify the inputs to the program tracking data base on program participation for the last month of the year.
- This one month gap means that it takes until February 1 for the statewide evaluator, Navigant, to even pull its sample of participants for the entire year to estimate program savings.
- After the sample is pulled, it can take one to two months to analyze the data from the tracking data base, customer bills and site visits to produce estimates of savings at the program level. Then it usually takes one to two weeks to actually draft the report, circulate for reviewed by sponsors and ultimately produce a report.
- Finally, it will take the independent evaluator, Itron, from two weeks to one month to review the work presented by Navigant, perform its own independent quality checks on the data analysis and measure verification and produce an independent analysis of the energy savings produced by the programs.

In sum, it normally takes up to four months from the program end date to the production of an independent report, which also needs to be reviewed and potentially modified based on these comments. Consequently, alternatives to the current schedule, which called for production of the final savings report in May, two months after the legislative report deadline, needed to be developed.

Itron and Navigant representatives met and generated the following options to meet the need to develop verified estimates of the previous year's programs no later than February 1:

- Option 1- Change the program reporting period from a calendar year basis to a fiscal year basis with the start and end dates being June 1 and May 31, respectively.
- Option 2- Produce the initial draft of the statewide evaluation of program savings by December 1 of each year by reducing the scope of savings to be included in the first report to those customers who had participated in the programs by November 1 of each year.[1]
- Option 3- Reduce the number of programs to be evaluated by the close of the year and ask Itron and Navigant to work closely together on all aspects of the evaluation to reduce the time needed to independently verify savings at the end of the year.

We chose to recommend Option 2 because this approach has worked in other states facing similar deadlines near the end of the program year and allows for the evaluation of a broader number of programs in the first few years of program start-up. Option 2 also gives Itron six weeks to review Navigant's work and produce its own independent analysis by January 15 and a final draft by February 15. This is probably the best we can do without moving the March 1 date or agreeing to a staggered evaluation approach where some programs are not evaluated every year but wait for two years to have a completed evaluation.

Option 1 was not recommended because the lag between the end of the program year May 31 and the report to the legislature (9 months later) was considered too long and the EmPOWER MD legislation sets goals on a calendar year basis. This option would also cause significant transitional expenses for program administrators who have set up their reporting and tracking systems on a calendar year basis.

Option 3 was not selected for three main reasons:

- The time saved by reducing the number of programs to be evaluated would still not allow Navigant to produce a draft estimate of savings until January 15 at best and more likely a complete draft, given tracking system delays, would be finished by January 30.
- Given a draft date of January 15-30, Itron would still need at least two weeks to produce a draft of an independent analysis of the savings data, and another week or

[1] Estimating annual program savings on December 1st would require the statewide evaluator to forecast the likely levels of participation from November 1st to December 31st. These forecasted results could then be trued up by February 1st of each year by Navigant or the PSC in its March report to the legislature.

two for public review, yielding a final draft of program savings between February 15 and February 28, which is probably too late to be included in the March report.

- The opportunity costs of deciding to limit the number of programs that will be evaluated by an independent firm are likely to be very high, particularly since many of these programs are relatively new and untested. It is better to aim for complete coverage of all programs in the early phases of program rollout, even if some evaluations are simply desk top reviews of assumptions and or process evaluations

Given our recommendation to support Option 2, we map out the corresponding schedule, which is shown in Table 6-1 and Table 6-2 below.

6.3 Recommended Schedule

In this section we present the recommended schedules for evaluation planning, execution and reporting and discuss how they mesh with other related processes including submission of documents to PJM and providing information to feed the planning process for 2012 to 2014 programs. We present two schedules: Table 6-1, which is geared toward understanding the relationships between evaluation of program load impacts and program reporting, and Table 6-2, which is geared toward understanding the relationships between the completion of process and market evaluation and the program planning process being conducted by the Maryland PSC staff.

6.3.1 Evaluation Planning and Reporting

Table 6-1 presents the recommended schedule for the production of evaluation documents and the reporting of program savings results to the Maryland PSC and ultimately the Maryland legislature. The table contains a description, start date and end date for each activity and which organization is responsible for the production or completion of each item.

Table 6-1: Evaluation Reporting and Planning

Key Dates	Deliverables
June 14, 2010	Draft Strategic Evaluation Plan
Aug 15, 2010	Final Evaluation Plans (Navigant will stagger high and low priority programs)
Dec 1, 2010 and Jan 15, 2011	Draft and Final Statewide Evaluation Report of 2009-2010 Program Savings (Navigant)[2]
Jan 10, 2011 and Feb 15, 2011	Draft and Final Verification of 2009 and 2010 Statewide Program Savings Report (Itron)
Jan 30 annually	Utility Programmatic Savings Reports - Includes reported and verified savings for previous program year (e.g., 2010 programs for the Jan 30, 2011 report)
March 1, 2011	EmPOWER progress report to State General Assembly
March 15, 2011	Utilities submit cost-effectiveness analysis for major programs and the entire portfolio for program years 2009, 2010, and 2011 program years.
May 1, 2011	Post-Installation Report to PJM for program savings bid into the market[3]
May 1 annually	Final Process Evaluation Results and Recommended Design Changes – interim results to be provided to utilities throughout the year

Key items in this schedule include:

Final Statewide Evaluation Plans for 2009 and 2010 Programs - due August 15, 2010 – We expect that Navigant will produce the evaluation plans for the energy efficiency programs and each utility program administrator will produce evaluation plans for the demand response programs.

Draft and Final Statewide Evaluator Reports - due December 1, 2010 and January 15, 2011, respectively – Section 6.2 provided the rationale recommending this report be filed before the end of the program year. The rationale for filing an updated report on January 15, 2011 is to provide sufficient time for the statewide evaluator to respond to comments and update the forecast of fourth quarter participation levels with data on actual participation levels available after January 1, 2011. In addition, the two weeks of holidays make it likely that there will only be four full working weeks to get the report published.

[2] Note the January 15 estimate of verified savings for 2010 programs does not need to be filed with the Maryland PSC but should be completed and sent to each administrator and Itron by January 15 to allow each administrator sufficient time to include these savings estimates in their January 30 Annual Program Savings Reports .This will also give Itron sufficient time to complete its independent verification of these savings numbers by February 15 of each year.

[3] The May 1 date for a post-installation M&V report to PJM for 2010 is dependent on whether utilities choose to bid into the 2011/2012 PJM Capacity Market.

Draft and Final Verification of 2009 and 2010 Program Saving Report from the Independent PSC Evaluator - due January 7, 2011and February 15, 2011, respectively – This schedule gives Itron 5 weeks to review the draft statewide evaluation report, provide comments, and incorporate this information into a draft verification report by January 7, 2011. This may be difficult given that this report must include estimates for both energy efficiency and demand response programs run in 2009 and 2010. We propose a final independent evaluator report due date of February 15, 2011 to allow time to review the final statewide evaluation report and the January 30, 2011 filing of each utilities' programmatic savings which may include savings from programs not included in the statewide evaluator report.

Utility Programmatic Savings Reports - due January 30 of each year – This report will provide the most comprehensive estimate of savings from energy efficiency and demand response programs based on ex ante savings and the latest participation counts from the tracking systems. It may also include preliminary estimates of the cost effectiveness of the program portfolios.

Maryland PSC report on EmPOWER Maryland progress to the State Assembly - due March 1, 2011 – This report can take advantage of three deliverables presented above: 1) the statewide evaluator report (with verified savings for some or most programs), the utility program reports due January 30, 2011 and the independent evaluation report due February 15, 2011.

Cost-effectiveness estimates – due March 15, 2011 – These estimates rely on the statewide evaluation results and will take a few weeks to finalize after the statewide evaluation results are finalized. Note that the Commission may determine that the alternative option of filing a minimal update on February 15, 2011 and a comprehensive update by May 15, 2011is necessary.

Post Installation Report to PJM – due May 1, 2011 – The actual deadline is May 15, but we recommend the report be completed two weeks early to be conservative. The purpose of this report is to provide evidence to PJM that each utility's previously nominated capacity from its EE programs has been verified using the evaluation plans previously submitted to PJM and any additional information generated by the program savings reports listed above from both the statewide and independent evaluators. The annual savings reports, due in January and February of each year, will give utility program administrators more than sufficient information to use in making this filing.

6.3.2 Integration of Evaluation Inputs into Program Planning Process

Table 6-2 provides report dates, along with a date for a proposed workshop to discuss proposed changes in program design suggested or supported by the process and market evaluation research. The schedule proposes that the statewide evaluation team shoot to produce the relevant process evaluation by October 1 of each year to provide useful information for program administrators and other parties to use in considering the potential need to modify program designs and or create new programs based on evaluation findings and other available information. Ideally this information would be made available by June 1 of each year to correspond to the current PSC staff schedule that starts on June 1, 2010 and culminates with the filing of program plans in March 2011. However, due to the late start of the evaluation effort and the potential need to modify current program approaches by January 1, 2011, we recommend that these process evaluations be made available by October 1 in this start up year.

Table 6-2: Integration of Evaluation Inputs into Program Planning Processes

Draft Schedule for Evaluation Planning, Execution, & Reporting in Maryland				
#	**Deliverable Description**	**Start date**	**End Date**	**Responsible Organization***
1	Recommend specific ex ante savings values and or algorithms from the Mid Atlantic States or other TRM's if viable to be included in utility planning data bases and a schedule/method for incorporating them into utility tracking data bases	5-Jun-10	15-Jan-11	INPSC & SWEVAL
2	Hold workshops/ meeting to gather comments and revise recommendations on draft report where appropriate. MD PSC staff/Itron publish final set of "adopted values" for incorporation in utility tracking databases by end of 4th quarter 2010	12-Jul-10	31-Dec-10	PSC
3	Process and Market Evaluations to be complete every May 1 to feed into June 1 program planning start date – Interim reports to be provided throughout the year.	30-Mar-10	1-May-10	SWEVAL &PA
4	PSC solicits comments on revised and new program designs based on evaluations and or any other market research		4-Oct-10	PSC
5	Recommended program design changes from Independent Evaluator, SW evaluator, and PSC staff are due		11-Oct-10	SWEVAL, PSC &INPSC
6	Workshop to Discuss Recommendations		18-Oct-10	PSC
7	PA reviews all comments and proposes program design changes and new programs based on process evaluations, internal input and mkt research from previous year		30-Jan-11	PA
*** Organizations KEY**				
INPSC= Independent Eval/Verification team		PSC- Public Service Commission Staff		
SWEVAL=Statewide Evaluator		PA= Program Administrators		

Appendix A

An Inventory and Assessment of Existing Utility Cost-Effectiveness Tools

A.1 Overview

Itron staff collected, and conducted a high-level assessment of, existing utility cost-effectiveness tools. The collection process included requests to utilities for all cost-effectiveness tools, review of all documents received from the PSC, and search of related dockets on the PSC website. The objective of this analysis is to: 1) identify actions that will be required by utilities to be able to calculate program cost effectiveness based on evaluation results anticipated in December of 2010, and 2) provide a schedule that would allow program cost-effectiveness estimates to be included in the March 2011 report to the Maryland General Assembly.

This section provides:

1) *Inventory of existing cost-effectiveness tools and future plans.* Including cost-effectiveness spreadsheets and calculators, method documentation, assumptions and data. What are utilities' plans regarding further development of cost-effectiveness tools? What are the PSCs plans as we understand them?
2) *Assessment of existing EE and DR cost-effectiveness tools.* What do utilities currently have in the way of cost-effectiveness tools? What are the capabilities and limitations of these tools? Do tools comply with existing PSC requirements and/or guidance? Who are the points of contact within each utility?
3) *Recommendations.* What cost-effectiveness capabilities and tools should the utilities be developing? When? Why?

A.2 Inventory of Existing Program Cost-Effectiveness Tools

The utilities employ separate cost-effectiveness tools for EE and DR. We have collected most of the available utilities DR cost-effectiveness tools. The focus of this discussion, however, is on EE cost-effectiveness tools.

As shown in Table A-1, all five utilities submitted EE program cost-effectiveness estimates and assumptions as part of their 2008 EmPOWER Program Plan filings to the PSC. There

was some apparent back-and-forth between the PSC regarding reporting of the PAC and/or the RIM tests, with the conclusion being that the RIM test is less applicable in MD since power is purchased through the PJM wholesale market and thus costs and benefits are spread across the entire PJM rate base. In the end, the RIM test generally was not reported, but is still calculated in the spreadsheets filed with the 2008 EmPOWER Program Plan filings. We could find no PAC test reported in the AP filing, though the requisite data is contained in the spreadsheets.

All of the EmPOWER Program filings contained, at a minimum, values-only flat file extracts from the EE cost-effectiveness spreadsheet tools. AP, PEPCO and DPL included the actual spreadsheet tools that they used for their cost-effectiveness calculations as part of their 2008 Program filings, though the PEPCO and DPL spreadsheet tools are not intended to be available to the public.

To date, Itron has obtained the complete EE spreadsheet tools (with formulas) for AP, BGE, DPL and PEPCO. We have outstanding requests for the spreadsheets with formulas from SMECO. PEPCO and DPL also have provided Itron with simple spreadsheet cost-effectiveness tools (using TRC test) that program implementers are using to screen EE measures and projects in real time. The other utilities are reportedly not using any cost-effectiveness screening tools as part of their program implementation.

Table A-1: EE Program Cost-Effectiveness Tools Inventory

Utility	Existing Tools	Cost Tests	Values or Formulas Submitted in Filings	Values or Formulas Received by Itron	Plans for Tool Development
AP	2008 EmPOWER Filing	No PAC	Values & Formulas	Values & Formulas	None
BGE	2008 EmPOWER Filing	All	Values & Formulas	Values & Formulas	None
DPL	2008 EmPOWER Filing & Program Screening Tool	All TRC	Values & Formulas NA	Values & Formulas Values & Formulas	None
PEPCO	2008 EmPOWER Filing & Program Screening Tool	All TRC	Values & Formulas NA	Values Only Values Only	None
SMECO	2008 EmPOWER Filing	All	Values Only	Values Only	None
Source: Cost-effectiveness spreadsheet tools provided to Itron by utilities.					

As shown in Table A-2, we have received DR cost-effectiveness tools (including values and formulas) from BGE, DPL, and PEPCO. AP has no DR programs. SMECO staff has promised the spreadsheets with formulas, but we have not received them as of the date of this memo.

Table A-2: DR Program Cost-Effectiveness Tools Inventory

Utility	Existing Tools	Program Cost Tests	Values or Formulas Submitted in Filings	Values or Formulas Received by Itron	Plans for Tool Development
AP	No DR Programs	NA	NA	NA	NA
BGE	2007 DRI Filing	B/C & Bill Impacts	Values & Formulas	Values & Formulas	None
DPL	2008 EmPOWER Filing	SC, TRC, PC, RIM	Values & Formulas	Values & Formulas	None
PEPCO	2008 EmPOWER Filing	SC, TRC, PC, RIM	Values & Formulas	Values & Formulas	None
SMECO	DRI Filing	PC, RIM, TRC, PAC	Values & Formulas	Values & Formulas	None
Source: Cost-effectiveness spreadsheet tools provided to Itron by utilities.					

To date, there has been no effort to link existing cost-effectiveness tools with reported or verified EE or DR program savings and none of the utilities have any current plans for further development of cost-effectiveness tools as part of the statewide evaluation.

A.3 Assessment of Existing Tools

The utilities' cost-effectiveness calculations generally appear to include the algorithms and assumptions necessary to calculate cost effectiveness, though there is some variation in tools and algorithms among utilities. As shown in Table A-3, the cost-effectiveness calculation tools provided by four of the utilities – BGE, DPL, PEPCO and SMECO – were developed by ICF and thus nearly identical in terms of algorithms and overall architecture.

AP developed its cost-effectiveness tools internally. While well-organized and seemingly user friendly, the calculations of individual measures are not entirely uniform, thus the AP calculation spreadsheets may be more difficult to integrate with the statewide evaluation reporting system than the ICF tools. The ICF tools also include some options that are not in the AP tools, such as application load shapes.

A.3.1 Economic assumptions

Some of the general assumptions vary significantly among utilities, including even among the ICF calculations. For example, AP, BGE, and SMECO use real utility and societal discount rates that are 3 percentage points higher than the ones used for PEPCO and DPL. [1] Additionally, SMECO's environmental adder is twice that of the other utilities.

Table A-3: EmPOWER Program General Assumptions

Utility	Utility Discount Rate (%)	Societal Discount Rate (%)	Inflation Rate (%)	Incentive Basis (%)	Societal Adder (cents)
AP	8.9	4	2.3	50% incremental cost	1 (carbon)
BGE	8.5	5	2.5	Net Payback years C = 1.5, R= 1	1.115
DPL	5.5	3	2.5	Net Payback years C = 1.5, R= 1	1.115
PEPCO	5.5	3	2.5	Net Payback years C = 1.5, R= 1	1.115
SMECO	8.62	3.03	2.5	Net Payback years C = 1.5, R= 1	2

We have not closely examined the measure-level assumptions and algorithms. At a glance, most of the measure specific input values seem reasonable, though there is some lack of documentation of the underlying assumptions, some ex ante values need to be revisited (e.g., incandescent exit signs as baseline), and the TRCs for some measures seem very high. Many of AP's measure calculations rely on EPA ENERGY STAR calculators and the sources appear to be somewhat dated. In the ICF-generated spreadsheets, for example, the documentation for many assumptions and algorithms is simply "ICF" and we are unable to find any documentation for the avoided cost assumptions, including even the year they were developed.

As shown in Table A-4, BGE, DPL and PEPCO include measure load shapes to estimate peak impacts within their calculations. The fact that AP has no peak load reduction programs planned (as shown in Table 2) suggests AP has no current capacity challenges, thus the load shape adjustments may not be viewed as useful. It is less clear why SMECO, which uses the ICF tool, has chosen to exclude this input variable.

Net-to-gross assumptions are included in all of the utilities' cost-effectiveness calculations at the program level. The NTG assumptions are not particularly well-documented, nor are they consistent among utilities, even the ICF utilities. SMECO and AP appear to provide the

[1] All of the discount rates appear to be in real dollar terms.

greatest level of granularity in terms of distinguishing between net-to-gross ratios for different program types.

Various levels of detailed narrative discussion of cost-effectiveness calculations are contained in all of the 2008 Empower Program Plan filings, with SMECO providing the least detail. Four of the utilities – AP, BGE, DPL, and PEPCO – state clearly that its methodology is consistent with the *California Standard Practice Manual: Economic Analysis of Demand-Side Programs and Projects*.[2] The SMECO methodology appears at a glance to be consistent with the *California Standard Practice Manual*, but there is no mention of it in the utility's EmPOWER Program Plan filing.

All five utilities provide multiple cost-effectiveness tests at the program level. BGE, DPL and PEPCO provided measure-level TRC estimates. AP reports four different cost-effectiveness tests results at the measure level, namely the TRC, PC, SC and RIM tests. SMECO does not report cost-effectiveness for individual measures.

Table A-4: EmPOWER Programs General Assumptions and Level of Detail

Utility	Measure Load Shapes	NTG (%)	Measure Level Detail	Measure Level Documentation	Measure Level Cost Tests	California Manual Cited	Tool Developer
AP	No	89-92 Res, 94-100 Com	Medium	Medium	TRC, RIM, PC, SC	Yes	Internal
BGE	Yes	90 large I&C,	High	Medium-Low	TRC	Yes	ICF
DPL	Yes	80	High	Medium-Low	TRC	Yes	ICF
PEPCO	Yes	80, except 100% low-income and DR	High	Medium-Low	TRC	Yes	ICF
SMECO	No	70-100	Medium-High	Medium-Low	TRC, RIM, PC, SC	No	ICF

Source: Cost-effectiveness spreadsheet tools provided to Itron by utilities.

Finally, there are many methodologies in the cost-effectiveness calculators that still need to be examined. For example, we have not reviewed the incentive cost allocation formulations for the programs, which could impact the cost-effectiveness test results; we are not aware of

[2] BGE cites two differences from the *California Standard Practice Manual*, namely that the *Manual* does not specify a societal discount rate (BGE uses 5%) or environmental adder (BGE uses 1.115 cents following PEPCO). It is not clear how these are variances from the *Manual*.

any clear guidance from the PSC on this issue and thus suspect there could be considerable amount of variation among utilities and programs within the utilities.

We also have not examined the process used for bundling of measures into programs. For example, BGE provides the following formulation for its large Industrial and Commercial programs (excerpt from 2008 Empower Program Plan):

> *The cost-effectiveness analysis was an iterative process, and included measure, program, and portfolio screening. First, all measures were screened using the measure incremental cost, (i.e., comparing an efficient appliance with a standard appliance). Those that had a benefit-cost ratio greater than or equal to 1.0 were included in a measure bundle that formed the basis for portfolio design. In some cases, measures that did not screen as cost-effective were included in programs. For example, each measure was screened by building type, to account for building energy usage variability. In many cases, a unique measure might be cost-effective in 6 out of 10 building types. From a program design perspective, it is not feasible to exclude certain building types from participation in a program offering that measure. Therefore, if a measure screened as cost-effective in building types and other configurations that could sustain a program, those measures in all building types were included in the program. Once the measure bundling was complete, the programs were screened via the TRC (non-incentive program costs). The portfolio was also screened using portfolio administration and communication costs.*

Whether other utilities are using a similar bundling approach, or have even explained their approach, is a topic for further investigation.

Appendix B

Demand Response Program Review

The Maryland utilities are investing substantial resources to solicit participants in demand response (DR) programs whose objective is to reduce peak demand on key high-load or high-cost days. Those peak demand impacts reduce the utilities' need for expensive generation capacity and/or purchases from PJM capacity markets, thus reducing costs. They also contribute toward the peak demand savings goals of EmPOWER Maryland and are bid into PJM's capacity market. The Maryland Commission is interested in documenting and validating the achievements of the utilities' DR programs in order to justify the program expenditures and ensure progress toward EmPOWER Maryland goals. The projections of the load reduction capabilities of the DR resources must also meet PJM monitoring and evaluation requirements in order to be accepted in the PJM capacity market.

B.1 DR Program Types

DR programs may be classified into various types. One useful classification is to distinguish between two approaches to achieving DR load reductions: 1) sending *price* signals to consumers, based on wholesale market prices, to which they may respond as they choose; and 2) sending *quantity* (curtailment) signals to customers, which they must obey under threat of financial penalty (because the consumers have typically been paid in advance for the right to curtail their usage under certain high-cost or emergency conditions).

Alternative mechanisms may be used to provide those price signals or quantity signals to consumers. For example, *price signals* may be sent through:

- Retail *dynamic pricing*; or
- *Economic* DR programs, in which consumers bid load reductions at specified prices, and receive payments for reducing load relative to a calculated baseline level of consumption if the bid is accepted.

Similarly, *quantity signals* may be sent through:

- A traditional *direct load control* (DLC) or interruptible service program operated by a distribution utility; or

- An energy service provider (ESP), curtailment service provider (CSP), or ISO, for an *emergency or capacity-based* DR program.

Currently, the DR programs offered by the Maryland utilities are largely of the second type, including air conditioner DLC and legacy interruptible service programs. For example, Baltimore Gas and Electric has a mix of residential customers with legacy DLC switches and a rapidly expanding number of PeakRewards customers with either new air conditioner switches or thermostat controllers.

B.2 Evaluation of Program Performance

Providing confidence that DR programs can, have, and will perform to expectations requires certain types of analysis, or evaluation. Perhaps most important is analysis of the performance of DR programs during actual events or test events.[3] For example, PJM has developed requirements for producing estimates of load reductions for various types of DR resources, such as DLC, Firm Service Level, and Guaranteed Load Drop. Particularly relevant to the Maryland utilities are PJM guidelines for reporting estimated per-participant load impacts of DLC programs. The guidelines have two parts, covering *estimation of average per-customer impacts* for active program participants, and *switch operability studies* that are designed to determine the percentage of active switches that are operable.

Utilities or curtailment service providers who operate DLC programs may choose to conduct load research studies to estimate load impacts, or, for programs that use a radio signal, they may opt to accept the tables of load impacts contained in the PJM-sponsored "Deemed Savings Estimates for Legacy Air Conditioning and Water Heating Direct Load Control Programs in PJM Region" report.

In the future, if the utilities and Commission decide to install smart meters and adopt some form of dynamic pricing (such as critical-peak pricing), then different types of load impact evaluations, which take advantage of the availability of hourly interval load data, will become appropriate. These include, for example, regression analysis of hourly data for the summer months, including event days (*e.g.*, days on which CPP prices apply) and non-event days, with the objective of estimating the magnitude of load reductions on event days.

[3] In the California demand response protocols, this type of analysis is referred to as *ex-post* load impact evaluation.

Appendix C

BGE Commercial Prescriptive Program: Illustrative Program Theory and Logic Model

C.1 Introduction

This section presents a preliminary analysis of the logic and underlying theory for BGE's Commercial Prescriptive Program (Program). First, we present the rationale for developing a program theory and logic model. Second, we present a brief description of the Program based on BGE's 2008 EmPOWER Maryland Program Plan, identified market barriers, the energy and demand goals, external factors that might influence program design and implementation, and the relationship of the Program to other programs. Third, we provide a logic model that illustrates a set of interrelated Program activities that combine to produce a variety of outputs that in turn lead to key short-, mid- and long-term outcomes. For each of the logic model elements, we suggest indicators that could be used to track performance. Finally, we assess the underlying theory or rationale for the Program design, relating it to accepted industry best practices.

To develop the Program theory and logic model and assessment, we relied on BGE's 2008 EmPOWER Maryland Plan and 2009 EmPOWER Maryland Annual Report, along with various secondary literature specifically related to the logic and design of programs similar to BGE's Commercial Prescriptive Program. This document is intended as a discussion document for BGE program implementation and evaluation staff. The final document – to be developed in conjunction with EmPOWER Maryland utilities and the statewide evaluator – will serve as a logic model discussion of the BGE Commercial Prescriptive Program and, more importantly, as a template for development of logic models for other key programs by all the EmPOWER Maryland utilities.

Importantly, program design and administration, along with the logic and theory behind it, is a dynamic process that should be responsive to changes in market conditions. In short, this and other program logic discussions should not be written in stone, but rather they should be treated as living documents that can and should be revised periodically as conditions change.

C.2 Motivation

An evaluation that is driven by the underlying theory of why program activities are expected to create specified outputs and outcomes is clear, systematic, and can lead to cost-effective determination of program effectiveness. According to Chen:

> ....specifying the underlying theory of a program within the evaluation allows that theory to be tested in a way that reveals whether program failure results from implementation failure or theory failure. Program theory clarifies the connections between a program's operations and its effects, and thus helps the evaluator to find either positive or negative effects that otherwise might not be anticipated. It also can be used to specify intermediate effects of a program that might become evident and measurable before final outcomes can be manifested, which can provide opportunities for early program assessment in time for corrective action by program implementers. (p. 29)

A theory-driven approach is particularly useful when a new program approach is being undertaken since many of the cause and effect relationships might be untested and implementation problems are likely to be more numerous. Program theory and logic models can be used to develop performance indicators that can be monitored over time and reported. Regularly providing such information to program managers can allow them to make mid-course correction in the design and/or delivery of the program activities. The program theories and logic models will also be used to identify high priority research opportunities, such as the effectiveness of key program elements or additional market research, which can further inform the design and delivery of the program activities.

Additional information on program theory and logic models can be found in Rogers, Hacsi, Petrosino, and Huebner (2000), The TecMarket Works Team (2004), Frechtling (2007) and Knowlton, Wyatt and Phillips et al (2009).

C.3 Program Description[4]

BGE's Large Commercial Prescriptive Program offers financial incentives to large C&I customers, relying heavily on equipment vendors and installation contractors to promote the program offerings and ensure lasting market transformation. The Prescriptive Program is one of a complement of programs targeting the commercial sector including a custom

[4] This section is entirely based on *BGE's Proposed Programs for Meeting EmPOWER Maryland Goals*, September 29, 2008 and includes numerous verbatim excerpts. As utilities and the statewide evaluator work together to develop logic models for individual programs, descriptive detail contained in existing utility program plans and other documents should be used to the extent possible to avoid unnecessary effort.

program and retro commissioning program. The large Commercial Program targets customers with a monthly billing demand of 60 kW, which includes all mid-size and large commercial and industrial customers and represents a very diverse range of customers in terms of size and type.

The Large Commercial program represents about half of total commercial program offerings (as measured by projected kWh savings).

C.4 Program Goals

To realize EmPOWER Program plan goals for 2011, BGE estimated that 665 customers would install 181 different types of measures achieving:

- 254 GWh energy savings and
- 41 MW reduction in peak demand.

The only evidence cited by BGE that these goals are reasonable was based on an assessment of savings potential from two sources. One was a previous experience with commercial customers while the second was a 2008 study by the American Council for an Energy-Efficient Economy. More information is needed on both of these sources.

C.5 Market Barriers

A review of BGE documents identified five barriers to energy efficiency improvements in the targeted large commercial sector that the Program is designed to overcome, including:

1. Low levels of awareness due to a lack of energy efficiency information
2. Lack of easy access to qualified vendors and installers,
3. Absence of tools to quantify savings,
4. The need to overcome perceived higher "first costs," and
5. Lack of access to capital.

The first three of these barriers can be mapped into the market barriers identified by Eto, Prahl and Schlegel (1996). For example, barrier #1 is consistent with *information or search costs,* barrier #2 is consistent with *hassle or transaction costs,* barrier #3 is consistent with *performance uncertainty*. However, Eto, Prahl and Schlegel do not consider first costs (barrier #5) to be a market barrier although incentives are a primary program strategy to increase adoptions.

> High first cost arises naturally in DSM program; many are designed to increase market adoption rates for energy-efficient products or services by reducing first cost (for example through rebates or other forms of financial assistance). We think that there is a basic difference between market barriers and the strategy used to overcome them. Thus, while reducing first cost may be an effective strategy to increase market adoption, we do not consider high first cost to be a market barrier, which this strategy has overcome. (p. 12)

While barrier #5 appears to be similar to *access to financing* barrier identified by Eto, Prahl and Schlegel, there is no financing component to the Program. It could be that access to capital is addressed through rebates alone.

As we will see in more detail later, the program is specifically designed to address these barriers. BGE documents state by raising awareness about energy saving opportunities and increasing access to qualified vendors and installers. The program further eliminates the need for customers to have tools to quantify savings and it reduces the actual (as opposed to just the "perceived") first cost of the measures. By reducing the first cost of the prescriptive measures, it also may help to address the lack of access to capital. We will also see that various elements of the program appear to be addressing additional market barriers identified by Eto, Prahl and Schlegel.

C.6 Key Program Elements

There are five key program elements which are designed to overcome the identified market barriers in order to meet the energy and demand goals.

1. ***Incentives.*** Incentives are available for predefined measures and eligibility requirements. The Program offers a simplified method to make efficient choices on pre-defined energy efficiency measures without requiring complex analysis or participation rules. Incentives and claimed savings are based on pre-defined technologies and calculation methods. Participants can choose from a menu of incentives for a wide range of pre-defined measures such as: lighting and controls, motors, variable speed drives, HVAC equipment, refrigeration equipment, office equipment, and food service equipment.

   BGE acknowledges that application and review process may need to be revisited based on feedback from customers. Moreover, to ensure cost-effective delivery of program savings, prescriptive rebate levels must be developed and updated periodically as market conditions change. Rebate levels that are too low will result in low participation; rebate levels that are too high will incur unnecessary costs on the utility and non-participant ratepayers. The BGE Large I&C Prescriptive Program

rebates are set at levels that are intended to give a net (after rebate) payback to participants of 1.5 years. If measure first-costs, energy prices, financing costs, or other cost-effectiveness factors change, the rebate levels will need to be revised. Calculation of rebate levels should include federal and state tax and other incentives.

2. ***Outreach.*** The program relies heavily on BGE account representatives and trade allies to increase awareness about the program, along with bill inserts, websites, and other forms of outreach such as seminars and conferences. Building strong relationships with trade allies (equipment vendors and installation contractors) is a key strategy for promoting prescriptive incentive availability to customers.
3. ***Training and Education of Trade Allies.*** Trade allies play a central role trade allies play in BGE's *Large I&C Program. Program administrators rely on them* to increase consumer awareness of the program and EE opportunities, selling and installing the eligible measures, and, over the longer term, continuing to work with customers to find additional energy savings opportunities. They are expected to play a key role in transforming the market. Given their importance, they will receive more detailed and comprehensive training. Trade allies may need a variety of education assistance, including (for example) program eligibility requirements and application processes, correct installation of and commissioning of measures, key benefits (selling points) of measures, and offerings of other BGE and government incentives and technical assistance.
4. ***Coordination with Other Programs.*** Numerous other programs may be applicable to BGE's large I&C customers, including BGE's large I&C audit and recommissioning programs. In addition, over the last several years a plethora of state and federal programs and tax incentives have come available, including (for example) substantial federal tax deduction for commercial building HVAC and lighting improvements. BGE account executives and trade allies need to be aware, and make customers aware, of these offerings.
5. ***Customer Feedback.*** A final program element worth mentioning is informing customers about effects of their participation in the Program. Often customers who participate in an energy efficiency program are unaware of the benefits of their participation (e.g., lower electricity bill). As a result, the positive benefits of the Program cannot serve to reinforce the desired behavior (the purchase of energy efficient equipment) thus reducing the chances that the desired behavior will be sustained. The solution is to make each participant more aware of the benefits of the Program (e.g., lower usage, lower bills, reduced greenhouse emissions, etc.) through various means (e.g., e-mail messages, regular US mail, etc.).

Figure C-1 provides a preliminary mapping of the five program components to the five identified barriers as well as additional ones listed by Eto, Prahl and Schelgel (1996).

Figure C-1: Market Barriers, by Program Elements

	Program Elements				
Market Barriers	**Incentives**	**Outreach**	**Training and education of trade allies**	**Coordination with other programs**	**Customer Feedback**
Low levels of awareness due to a lack of energy efficiency information		X	X	X	X
Lack of easy access to qualified vendors and installers			X		
Absence of tools to quantify savings	X		X		
The need to overcome perceived higher first costs	X	X			
Lack of access to capital.	X				
Organizaton practices			X		
Asymmetric information		X			
Product or service availability			X	X	

* Eto, Prahl, Schlegel (1996) market barriers are highlighted.

Finally, there are a number of issues and questions that must be addressed before the program theory and logic model are complete.

- With respect to the outreach and training and education components, there are a number of questions. For example, the BGE documents contained no information about the number of equipment vendors and installation contractors who operate in BGE service territory. What percent of these do they plan to reach annually? In addition, what topics will be covered and in what depth? Such questions must be answered in order to assess the efficacy of the Program (i.e., can it reasonably be expected to produce the planned outputs and outcomes).
- With respect to coordination with other programs, more information is needed. For example, what activities will be carried out and at what frequency? It is also important to track the extent to which BGE participants access these other programs and for which specific measures. Using this information, the program incentive levels can be adjusted in future program years for specific measures or even end support for specific measures if support from these other programs appears adequate.

C.7 External Influences

In addition to energy prices, economic growth, inflation, there are a variety of other external influences that could hinder or support program participation and impacts, including:[5]

- Federal appliance and equipment standards
- Federal Commercial buildings tax deduction
- ENERGY STAR Program
- National commercial buildings initiative
- Maryland tax incentives

One key question for evaluators is whether these other programs will compete with or complement the BGE program. Will customers participate in the most attractive program and, because of the hassle factor, not participate in others?

It is not clear from existing BGE documents that these programs have been inventoried and considered in the design of the Large Commercial Prescriptive Program.

C.8 Relationship to Other BGE Programs

BGE does not believe that there are any BGE proposed programs with the potential to create barriers to the success of other programs. BGE, working with its partners, have developed comprehensive programs that will both have an immediate impact on energy use and in the long run will help transform the market into one where customers seek energy efficient options on a regular basis no matter the incentives. This, of course, is another researchable issue for an evaluator.

C.9 Program Logic Model

The Program logic model, illustrated in Figure C-1, flows from top to bottom and is organized according to five basic categories:

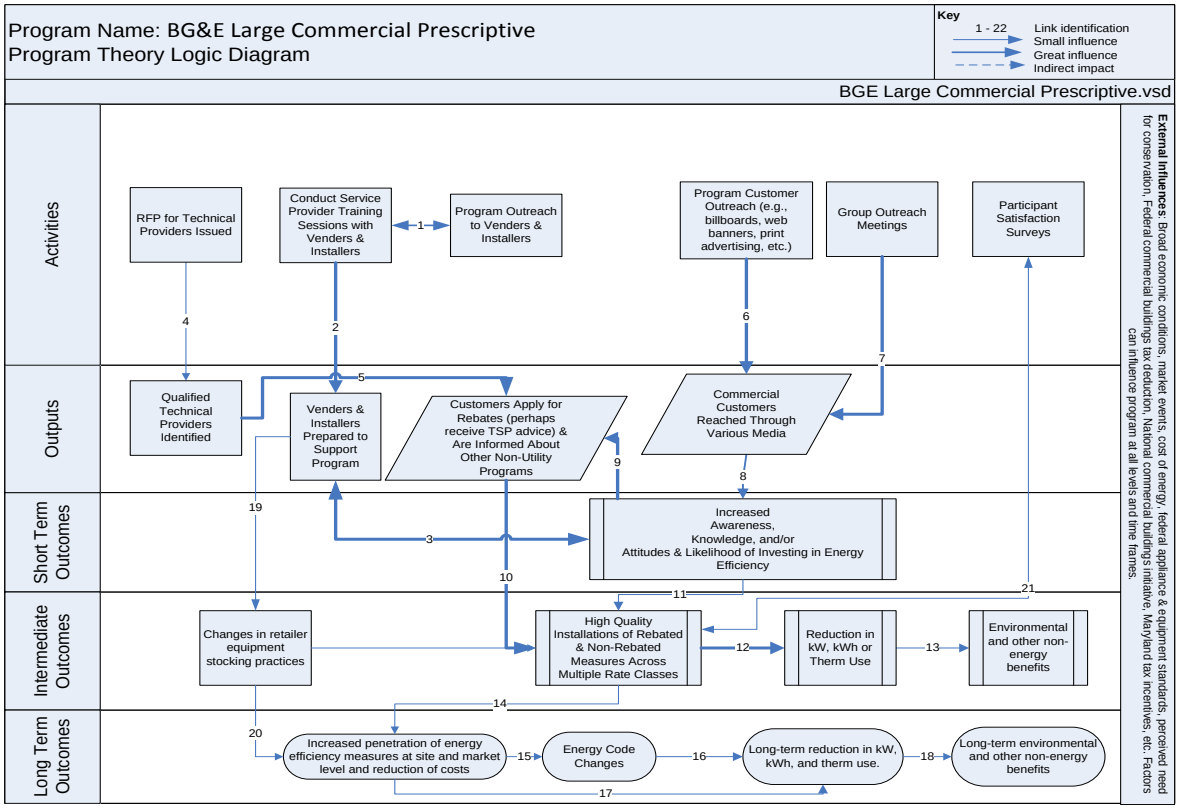
- Program activities
- Outputs
- Short-term outcomes

[5] Only some factors are listed below. This list will grow over time as more external influences emerge and are identified.

- Intermediate outcomes
- Long-term outcomes

Each link in the logic model is numbered. These numbers are to facilitate reference with the indicators in Table C-1. The numbers do not indicate a sequence of activities. The potential performance indicators for each link in this logic model are presented in the following section.

Figure C-2: Logic Model: BGE Large Prescriptive Commercial Program



Note: Making each participant more aware of the benefits of the Program (e.g., lower usage, lower bills, reduced greenhouse emissions, etc.) through various means (e.g., e-mail messages, regular US mail, etc.), while an important Program element, is not reflected in the logic model.

C.10 Performance Indicators

The performance indicators for each linkage in the logic model have been identified and are presented in Table C-1. These indicators characterized as "potential" since there is rarely the money or need to measure every indicator. The most important links must be identified and, for each selected link, the most important indicators for which data can reasonably be

collected must be also identified *and* carefully operationalized. Responsibilities for collecting these data (e.g., BGE or the evaluator) must also be determined.

Table C-1: Program Theory and Potential Indicators, by Logic Model Link

Link	Program Theory	Potential Indicators
1	Workshops held for service providers (venders and installers) to provide a general overview of the I&C portfolio of programs and to teach attendees how to participate in the programs and promote them to their customers.	Workshops organized according to adult education principles Number of seminars & workshops, by topic Number of participants Number of contact hours of instruction Stated intentions to promote Program to their customers Trainee satisfaction with curriculum and presentation
2	Workshops prepare service providers to support the program.	Increased awareness of the Program Knowledge of key Program elements (e.g., eligible measures, rebates and application process) Percent change of the number of active installers of select measures in 2012 compared to 2011
3	Properly trained service providers use their knowledge of the Program to sell & properly install efficient measures.	Degree to which trainees incorporate Program information into sales messages The percent of measures properly installed by participating venders. Numbers of customers influenced by service providers to participate in Program.
4	Once issued, the RFP for technical service providers (TSPs) generates a critical number of qualified respondents.	Number of firms responding Number of qualified firm responding
5	TSPs provide support to customers to insure quality installations	Percent of customer installations in which TSP provided added value TSP are able to solve installation issues as they arise Percent of customers who are satisfied with TSP services
6	Information via various media (e.g., billboards, web banners, print advertising, etc.) is delivered to commercial customers	Ease of website navigation Number of hits on BGE website Number of brochures/mailers delivered to target audiences Square feet of billboards displaying program information
7	Outreach meetings with key customer groups are held to inform customers about the Program.	Attendees as a percent of those invited Number of outreach meetings held Number of attendees at each meeting
8	Customers reached through various media and outreach meetings experience increases in awareness & knowledge, attitudes regarding the Program	Percent of customers recall receiving the message Percent of customers recall the source of the message Percent of customers recall details of the message Percent of customers understand the message Percent increased awareness and knowledge of the Program For select measures, percent of customers in 2011 that say they

Link	Program Theory	Potential Indicators
		trust reliability and/or quality of measures less percent of customers in 2012 that say they trust reliability and/or quality of measures. Percent increase in customers who say BGE is a trusted resource. Percent increase in positive attitudes change towards energy efficiency Percent increase in the self-reported likelihood of investing in energy efficiency
9	Increases in awareness, knowledge, and attitudes cause customers to apply for Program rebates and become aware of non-utility programs.	Number of customers who apply for an incentive Amount of incentives for each end use category Satisfaction with application process
10	Customers (across multiple rate classes) who receive Program rebates (and perhaps receive TSP advice) and are informed of other non-utility programs will have efficient equipment correctly installed.	Cycle time (time from application to receipt of payment) Percent of applicants meeting eligibility requirements Installation rate (percent of installations properly installed and capable of producing savings) Number of participants from various rate classes in 2009-2010 / Average number of customers from various rate classes in 2009-2010 Percent of customers who access the resources from non-utility programs.
11	Increases in awareness, knowledge, and attitudes cause customers to install rebated measures without a Program rebate, (i.e., non-participant spillover).	Installation rate (percent of installations properly installed and capable of producing savings) Percent of customers who access the resources from non-utility programs.
12	Properly installed measures produce expected (ex ante) energy and demand impacts. Increase in properly installed hardware and changes in behavior with respect to energy efficiency results in reduction in kW, kWh, and therms (including both participant and nonparticipant spillover)	Ex post gross realization rates Ex post estimates of gross and net energy and demand impacts, participant spillover, and nonparticipant spillover For each measure, reported number of participants in 2009-10 x reported technology units in 2009-2010 / forecast number of participants in 2009-2010 x forecast technology units in 2009-2010 (note: technology units are defined in the 2008 EmPOWER Program Plan) Percent of energy and demand goals achieved
13	Gross energy and demand impacts in turn produce environmental and other non-energy benefits	CO2, NOx, & SOx reduced Increased comfort, productivity, etc.
14	Over time, enough customers reduce energy such that increases in market share and penetration can be seen at the site and market level.	Percent increases in market share & penetration. Number of select measures sold in 2011 / Number of select measures sold in 2010.
15	Increased penetration of energy efficiency measures at site and market level eventually leads to	Number and types of code changes

Link	Program Theory	Potential Indicators
	changes in energy efficiency codes.	
16	Energy Code Changes lead to sustained reduction in kW, kWh, and therm use.	Long-term reduction in energy and demand.
17	Increased penetration of energy efficiency measures at site and market level leads to long-term reduction of kW, kWh and therm use.	Long-term energy and demand impacts.
18	Long-term reduction in energy and demand lead to long-term environmental and other non-energy benefits.	Long-term reduction in CO2, NOx Sox Increased comfort, productivity, etc.
19	Demand for more efficient measures by service providers leads to changes in equipment stocking practices	Number of stores stocking select measures in 2012 / Number of stores stocking select measures in 2011. Square feet of shelf space occupied by select measures in 2011 / Total shelf space for application in 2011.
20	Increase in stock of efficient equipment leads to increased penetration of energy efficiency measures at site and market level & reduction in costs	Percent increases in market share & penetration. Average unit price in 2009 dollars of measure in 2011 / Average unit price of measure in 2010 in 2009 dollars
21	Participant satisfaction surveys are conducted and results are provided to Program managers.	Percent of surveyed customers who report being very satisfied with the rebated measure. Extent to which feedback to Program managers take negative feedback into consideration when modifying the Program.

C.11 Discussion of Comparative Program and Theory Research

The Program is modeled on a variety of nationally recognized programs and targets small commercial and large I&C populations though all market channels, across various price points. BGE's portfolios were developed using a model that maximizes the cost-effective savings of each program. The model develops measure installation estimates using a technique based on results of peer-reviewed academic studies on the diffusion of technologies, including work by Lawrence and Lawton, Packey, and others.[6]

Further, it will encourage energy efficiency improvements with compelling value propositions and easy participation opportunities. More specifically, BGE's C&I Prescriptive Program consists of program offerings and delivery mechanisms that have been developed from best practice approaches and adapted to the BGE market. In designing the program, the following "best practices" strategies are being applied:

- Start with proven program designs while tailoring approaches for targeted market segments;
- Solicit stakeholder input upfront in the detailed planning process to understand the local market;
- Begin with tested technologies with well-established energy savings performance, and supplement them for market segments as appropriate;
- Launch with simple programs and evolve the design and delivery channels over time;
- Use building codes and equipment standards as baseline assumptions, supplementing them with baseline studies where appropriate;
- Design technical review, proof of project installation, project file documentation, and quality assurance techniques into program designs to ensure credible and reliable program impacts;
- Retain design flexibility to enable BGE to adjust specific designs as dictated by customer response and evaluation results, and to rebalance the portfolio based on individual program performance and emerging opportunities; and
- Invest in education, training, and outreach to build a strong local infrastructure.

The process of setting incentives began with a review of the incentive levels offered by other utilities. Remaining incentive levels were calculated as the dollar amount required to reduce customer payback times to levels acceptable by customers, which is 1.5 years for commercial and industrial customers and one year for residential customers. This calculation was performed on a measure-by-measure basis and, as such, yielded incentive levels that vary between 25% and 75% of incremental cost.

BGE offers the following rationale for prescriptive incentives:

- They offer a simple, easy to use process that encourages participation;
- They allow vendors to use rebates to make the sale by showing customers the reduced installed cost;
- They cover the majority of common energy saving measures across most customer and end-use markets;
- They help transform supply chain markets by giving distributors, contractors, and other trade allies clear signals on what is preferred in the market. That is, it helps change equipment stocking practices, which is key to energy efficiency market transformation;
- They facilitate a quick-start approach that enables BGE to realize rapid energy and demand reduction results; and

- They are an effective way to reach mid-market I&C customers who tend to be underserved.

Before the program theory and logic model are completed, one must identify relevant best practices in the literature and assess the extent to which the program reflects those best practices.

C.12 References

Chen, Huey-Tsyh. (1990). *Theory-Driven Evaluations.* Sage Publications, Inc.

Joe Eto, Ralph Prahl, and Jeff Schlegel. (1996). *A Scoping Study on Energy-Efficiency Market Transformation by California Utility DSM Programs*. (Berkeley, CA: Ernest Orlando Lawrence Berkeley National Laboratory). LBNL-39059 UC-1322, 9.

Frechtling, Joy A. (2007). *Logic Modeling Methods in Program Evaluation*. San Francisco, CA: Jossey-Bass Publishers.

Knowlton, Lisa Wyatt and Cynthia C. Phillips. (2009). *The Logic Model Guidebook: Better strategies for great results*. Thousand Oaks, CA: Sage Publications, Inc.

Rogers, Patricia J., Timothy A. Hacsi, Anthony Petrosino, and Tracy A. Huebner (Eds.) (2000). *Program Theory in Evaluation: Challenges and Opportunities*. San Francisco, CA: Jossey-Bass Publishers.

The TecMarket Works Team. (2004). *The California Evaluation Framework.* Prepared for the Southern California Edison Company.