NYSERDA EXISTING FACILITIES PROGRAM

FINAL Evaluation, Measurement, and Verification Plan

December 11, 2009

I. Introduction

The detailed evaluation plan presented in this document builds upon prior evaluation activities conducted for the Existing Facilities Program (EFP). In developing this evaluation plan, NYSERDA has incorporated feedback provided by the Department of Public Service (DPS) and the EEPS Evaluation Advisory Group (EAG), and has worked closely with its team of independent evaluation contractors to select the most appropriate evaluation approaches based on the current design of the program. This plan was developed to conform to the DPS evaluation guidelines released on August 7th, 2008 and to provide the highest level of rigor possible within the available resources.

As the EFP works to meet its current SBC program goals, NYSERDA and its evaluation contractors will closely monitor aspects of that process such as participation levels, achievement of near-term goals, and other programmatic issues in order to adapt this plan, as needed, to provide the most relevant and useful evaluation. For example, adjustments may be needed to sample sizes or research issues if assumptions about the program do not develop as initially anticipated. As such, NYSERDA views this plan as a flexible, living document that will be updated, as necessary, with appropriate notice to DPS and other interested parties.

This evaluation plan was designed to constitute a comprehensive approach to assessing the entire EFP supported by SBC funding.

II. Summary of Goals, Cost and Schedule for Evaluation Activities

The overarching goals of NYSERDA's **New York Energy $mart**^{SM} and EEPS Program evaluation efforts are to: (1) conduct credible and transparent evaluations, and (2) provide NYSERDA program staff and managers, the New York State Public Service Commission (PSC), DPS staff, and other stakeholders with timely and unbiased information regarding program implementation. Specifically, the goals for the EFP evaluation include:

(1) Establish rigorous and defensible estimates of the energy savings that can be attributed to the program;

(2) Develop a comprehensive understanding of current and emerging markets (*e.g.*, market structure and market actors), especially vendors of interval meters and energy service companies (ESCOs);

(3) Provide baseline and background information required by NYSERDA to define and deliver programs to target markets;

(4) Track changes in the market over time with a specific focus on market indicators that are likely to be impacted by program offerings, focusing especially on:

a) Existing Facilities Program market share as measured by energy savings and peak demand reduction, segmented by utility territory and market sector, and eligible contractor base;

b) Existing energy efficiency/demand response (DR) service delivery channels (*i.e.*, energy service companies (ESCOs), curtailment service providers (CSPs), meter service providers (MSPs), load serving entities (LSEs), and/or other DR service providers) and most active contractors;

c) Current participation rates in New York Independent System Operator (NYISO) and utility DR program offerings and EFP interaction with these programs;

d) Current penetration of advanced metering infrastructure (AMI) and smart grid platforms; and

e) Other metrics as identified.

(5) Assess and document the effectiveness of program processes in providing appropriate services and quality assurance for different types and sizes of projects; make recommendations for program improvement, focusing specifically on practical, actionable recommendations

(6) Assess end user and ESCO/vendor experience with the program; and

(7) Examine and identify the reasons for end user and ESCO/vendor participation in the program to assess:

a) Awareness, knowledge and reasons for participation;

b) Barriers to participation; and

c) The decision making process for participation, as well as perceptions of the impact of external factors on participation (*e.g.*, financial climate, climate change, energy prices, etc.).

The Existing Facilities Program budget is approximately $93 million (4th Quarter 2008 through 2011). The proposed evaluation budget is $3.0 M, or 3% of program funding.[1] NYSERDA believes this level of funding for evaluation is justifiable and adequate to achieve a high level of confidence and precision related to program impacts as well as address key process and market evaluation issues. The primary factors supporting an evaluation budget of 3% for this program are: the expectation that a large majority of the total expected savings from the program will come from a relatively small percentage of the participating projects; and the fact that the overall population of industrial facilities in the State is small compared to commercial or residential markets. Evaluation budgets are detailed in Table 1.

[1] This evaluation budget includes only external contractor costs. Other overarching evaluation costs, including NYSERDA's internal evaluation management and statewide study costs, are additional; however, the total evaluation costs will not exceed 5% of program funding at the portfolio level.

Table 1. Existing Facilities Program Evaluation Schedule and Budget

Evaluation Element	Estimated Budget and Completion						% of Total Budget
	2009	2010	2011	2012	2013	Total	
Market Characterization & Assessment	--	$203,000	$184,000	--	--	$387,000a	17%
Impact Assessment	--	$1,804,200	--	--	--	$1,804,200b	79%
Process Evaluation	--	$94,000	--	--	--	$94,000c	4%
Total	**--**	**$2,101, 200**	**$184,000**	**--**	**--**	**$2,285,200**	**100%**

a. Primary data collection costs represent approximately 37% of the total proposed MCA evaluation budgets.

b. Telephone survey costs represent approximately 5% of the total proposed impact evaluation budgets.

c. Primary data collection costs include $16,500 for interviews and $15,000 for the end user survey.

III. Existing Facilities Program Description and Goals

The EFP promotes energy efficiency and demand management. This new program is a consolidation of two prior NYSERDA programs -- the Peak Load Management Program (PLMP) and the Enhanced Commercial and Industrial Performance Program[2] (ECIPP). Building upon the success of these two programs, the July 1, 2008 merger is a streamlined program offering that is less complicated, and more accessible to potential customers. EFP targets a range of customer sectors including commercial and industrial businesses, healthcare facilities, universities and colleges, State and local governments, and mission critical facilities such as data centers and communications facilities.

There are two types of EFP incentives - pre-qualified and performance-based:

- **Pre-qualified** electric incentives encourage customers working on small-sized energy projects and equipment replacement projects to purchase and install more energy efficient measures. Some of the electric measures available to qualifying customers are lighting, heating, ventilation and air conditioning (HVAC), chillers, motors, variable frequency, and interval meters.

- **Performance-based** incentives are for customers or ESCOs working on large-scale projects. The incentives are typically higher than those for pre-qualified projects, and the performance-based projects require an engineering analysis, and are potentially subject to measure and verification requirements. The various types of performance-based incentives are expected to result in the following impacts:

    - Electric efficiency incentives encourage the implementation of projects that deliver verifiable annual electric savings.

[2] The Enhanced Commercial/Industrial Performance Program combined two early programs – Commercial/Industrial Performance and Smart Equipment Choices.

- Combined heat and power (CHP) incentives contribute to the installation cost of clean, efficient, and commercially available CHP systems.

- Industrial and process efficiency incentives help offset the costs of projects focused on increasing productivity, and increasing electricity consumption on a per-unit of production basis.

- Demand response provides help with a portion of the cost for technology, such as load curtailment and shifting (LC/S) and distributed generation (DG), that enable facilities to participate in the NYISO demand response programs (which reduce electricity load in response to emergency and/or market-based price signals).

Energy demand reduction contributes to improvements of New York's energy system reliability and security, while helping businesses and industries to reduce operating costs. Allowing customers, ESCOs, and contractors access to multiple incentive strategies to support their energy projects will enable the New York ESCO community to continue to grow the market for energy efficiency in existing buildings, process equipment, and non-building efficiency measures.

Anticipated MWh savings for the Existing Facilities Program are shown in Table 2.

Table 2. Projected GWh Savings for Existing Facilities Program (Five-Year SBC3 Goal)

	2006-2011 (Five-Year Goal)
Electricity Savings (GWh)	427
Permanent Demand Reductions (MW)	110
Callable Demand Reductions (MW)	240

Source: NYSERDA, *System Benefits Charge Proposed Plan for* ***New York Energy $mart***^{SM} *Programs (2006 – 2011)* As Amended, March 2, 2006, Page 4.5, Table 4.1 and Page 4.9, Table 4.3.

IV. Logic Model/Theory

Figures 1 and 2 are logic models of the two component programs of the Existing Facilities Program, the Peak Load Management Program and the Enhanced Commercial and Industrial Performance Program. As the program evaluation efforts defined in this plan begin, a first step in the process will be to review the logic models and make updates as necessary (see discussion in Section V).

Logic modeling activities will occur early in the evaluation process after completion and approval of the Detailed Evaluation Plan. NYSERDA's evaluation contractors convene logic model "workshops" with program staff to discuss program inputs, activities, outputs, outcomes, external influences and other elements that need to be documented in the logic model. The evaluation contractors then document these discussions in a brief program theory/logic report, which includes a logic model diagram for the program.

NYSERDA will invite DPS Staff to participate in logic model workshops and review draft program theory/logic reports.

Figure 1 - Peak Load Management Program Logic Model

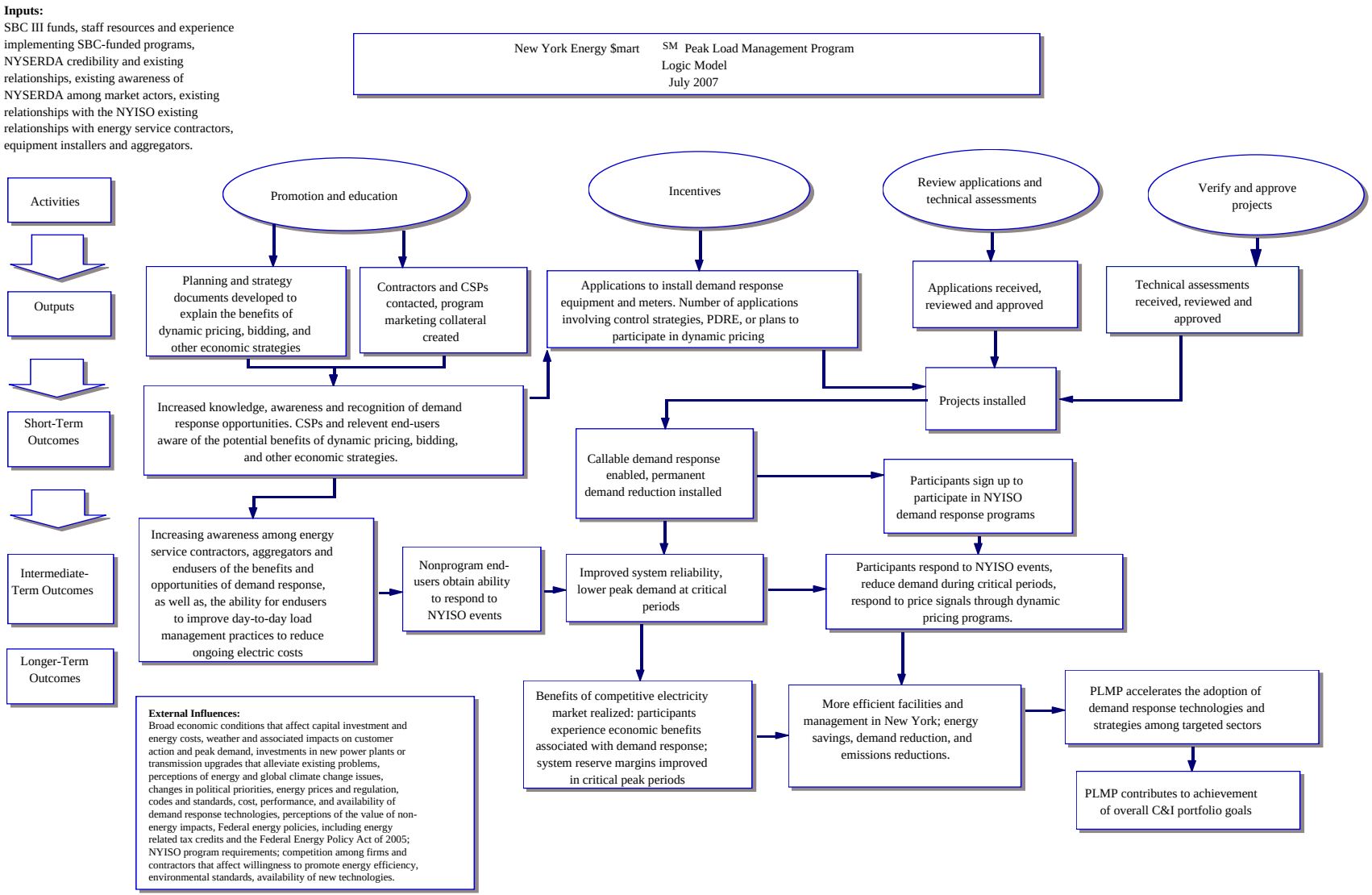
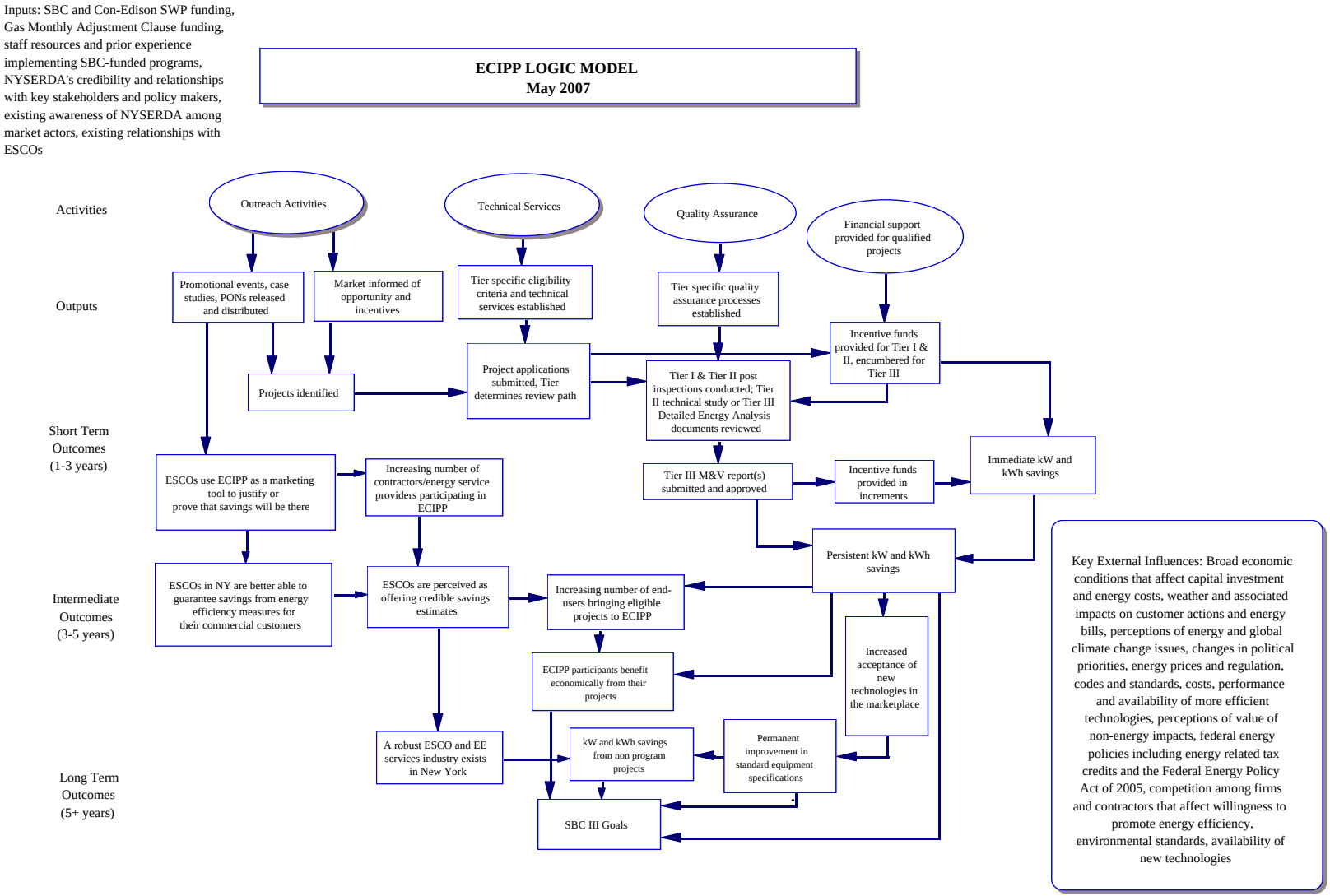


Figure 2 – Enhanced Commercial/Industrial Performance Program Logic Model



V. Market Characterization & Assessment Plan

This section presents the Market Characterization and Assessment (MCA) evaluation plan for the Existing Facilities Program (EFP). The plan is structured to accommodate the multiple components of the Program in that the MCA Team will focus its efforts on the demand response performance based incentive program component during the 2010 evaluation cycle then turn its attention to the energy efficiency performance based incentive program components during the 2011 evaluation cycle.[3]

Research Objectives

The primary goals of the MCA evaluation effort are: (1) to develop a comprehensive understanding of current and emerging markets (*e.g.*, market structure and market actors); (2) to provide baseline and background information required by NYSERDA to define and deliver programs to target markets; and (3) to track changes in markets over time with a specific focus on market indicators that are likely to be impacted by program offerings.

The proposed MCA evaluation plan was structured to accommodate these overarching research goals with a specific focus placed on the market and context within which the EFP operates. The plan was designed to validate program assumptions regarding market characteristics, provide additional details regarding market structure and opportunities, and ensure consistency with prior program evaluation activities conducted by NYSERDA. The continuity in approach will enable the MCA Team to build upon prior research findings and ensure that current and subsequent evaluation results can be used to assess progress towards meeting the PSC's public policy goals for the SBC programs, as well as the institutional goals NYSERDA has established to move markets towards improved energy efficiency. In addition, the evaluation results can be used by NYSERDA program staff and managers to adjust program implementation as needed to ensure maximum market interest and uptake of program offerings.

Activities

The proposed MCA evaluation plan for the EFP consists of multiple activities (blue arrows) and associated research tasks (bulleted lists), as shown in Figure 3. The approach is centered on the development of a Program Logic Model Report that will then be used to guide program-specific evaluation activities.[4] A variety of primary and secondary data sources will be used to generate information on a number of topics relevant to the program including: program accomplishments and market share in terms of both energy impacts and interaction with key market actor groups (*e.g.*, energy service companies (ESCOs), curtailment service providers (CSPs), demand response (DR) service providers, etc.); changes in end-use customer and contractor[5] awareness and understanding of the energy efficiency and peak demand reduction strategies promoted by the program; and changes in industry

---

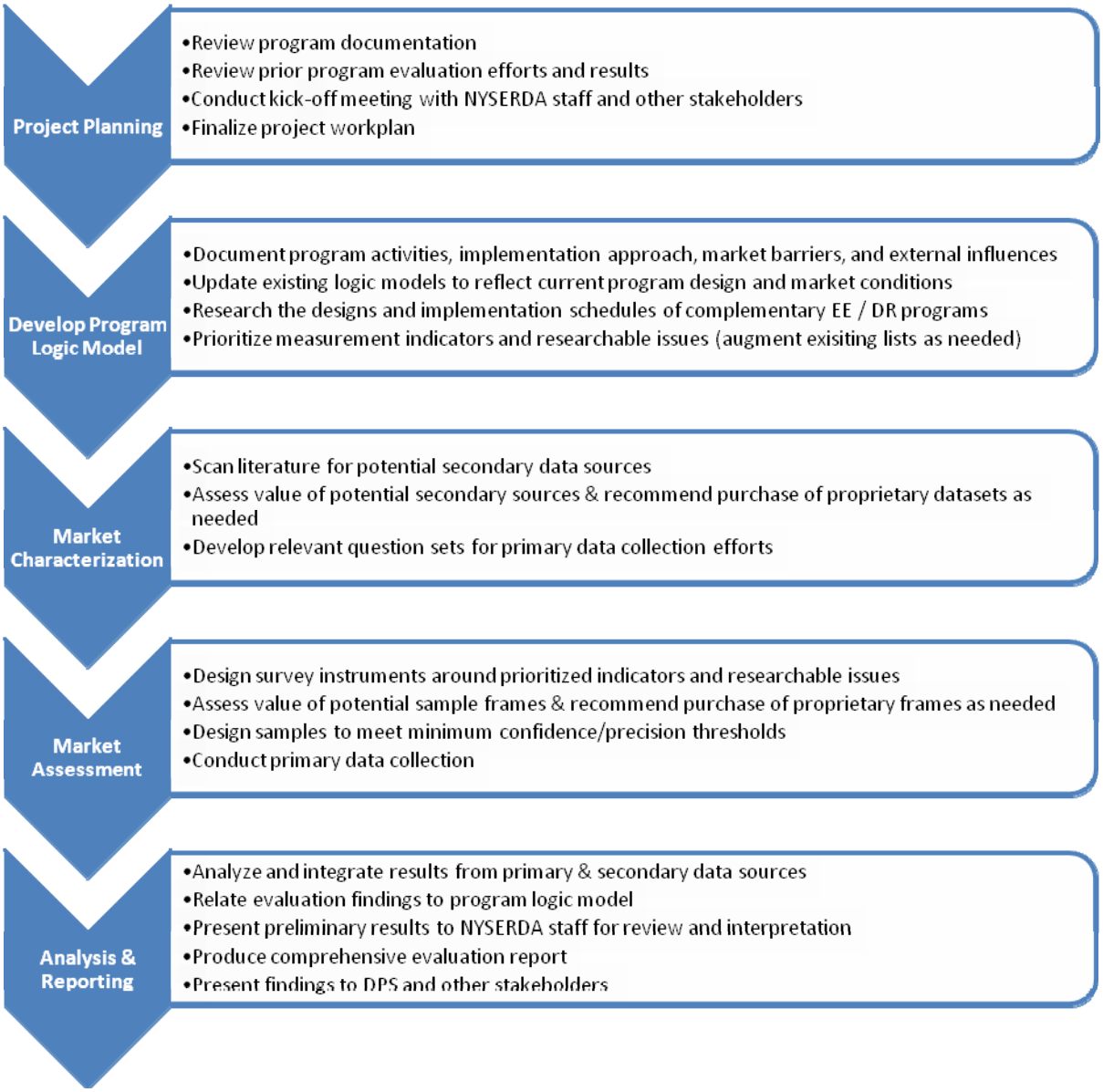
[3] The Industrial and Process Efficiency Performance Based Incentive Program component will be addressed during the comprehensive MCA evaluation of the Industry and Process Efficiency Program (see separate evaluation plan). Similarly, the Combined Heat and Power (CHP) Incentive Program component will be addressed during the comprehensive MCA evaluation of the DG/CHP Program (see separate evaluation plans).

[4] The Enhanced Commercial/Industrial Performance Program (ECIPP) and Peak Load Management Program (PLMP) were consolidated into the Existing Facilities Program. Program Theory and Logic Model Reports were developed in 2007 for the ECIPP and PLMP; however, a Program Theory/Logic Model Report has not been developed for the combined Existing Facilities Program.

[5] The term "contractor" in this workplan includes the following entities eligible to participate in the Existing Facilities Program: energy service companies (ESCOs), curtailment service providers (CSPs), meter service providers (MSPs), load serving entities (LSEs), and/or other demand response (DR) service providers.

infrastructure resulting from program interventions as measured by number of contractors active in the state, annual value of contractor activity, and market perceptions of industry strength and competitiveness. This information will be used to assess program execution across multiple dimensions, including the extent to which program interventions are helping to develop a viable contractor industry in New York, which is a stated objective of the program. Each activity and the associated research tasks are discussed in more detail in the remainder of this section.

Figure 3. Synopsis of MCA Evaluation Activities and Research Tasks



Project Planning

This task encompasses a variety of project planning activities including review of available program documentation and prior program evaluation results, meetings and discussions with NYSERDA evaluation staff and other evaluation contractors, a project kick-off meeting with EFP staff and other project stakeholders, and the development of the final project workplan. An important component of this initial phase of the project is providing EFP staff an opportunity to discuss research items of interest to ensure development of a research agenda geared toward overcoming any existing gaps in staff's knowledge of current market conditions and opportunities. The collaboration with NYSERDA program and evaluation staff and other project stakeholders will continue throughout the evaluation as iterative processes are used to review and finalize interim and final project deliverables (*e.g.*, survey instruments, summary memos and reports, etc.).

Develop Program Logic Model

The EFP is a new program; therefore, a Program Logic Model Report does not yet exist for the program.[6] Given the importance of the Logic Model Reports in guiding NYSERDA's program-specific evaluation activities, an initial activity undertaken by the MCA Team will be to develop a Program Logic Model Report for the overarching EFP. The Logic Model Report will: summarize the context within which the EFP operates; discuss the market barriers and inefficiencies that the program seeks to address; describe the program implementation approach and anticipated outputs and outcomes; develop a logic model diagram showing the linkages between program operation and anticipated outputs/outcomes; and identify relevant measurement indicators and researchable issues. As part of this exercise, the team will update the existing ECIPP and PLMP *Program Logic Model Reports*[7] to reflect current program designs and the state of the market. In addition, the MCA Team will research the designs and implementation schedules of complementary programs being administered by other entities[8] to identify potential leveraging opportunities wherein NYSERDA and the other program administrators can possibly collaborate to achieve broader and deeper program impacts.

Before proceeding, it should be noted that this initial phase of the evaluation will provide an opportunity for the MCA Team to generate feedback regarding proposed program design and implementation strategies. The Team will use the development of the logic model to suggest opportunities for program improvement, if any are observed, in the hopes of streamlining program delivery processes.

Market Characterization

Market characterization results will be generated primarily from secondary data sources, supplemented by information gathered during primary data collection efforts. Key data sources to be used for this activity include the Existing Facilities Program tracking database, previous program evaluation reports prepared for NYSERDA and for similar programs operating in other jurisdictions, McGraw-Hill Construction Dodge databases, U.S. DOE's Commercial Buildings Energy Consumption Survey (CBECS) data,

---

[6] The Enhanced Commercial/Industrial Performance Program (ECIPP) and Peak Load Management Program (PLMP) were consolidated into the Existing Facilities Program. Program Theory and Logic Model Reports were developed in 2007 for the ECIPP and PLMP; however, a Program Theory and Logic Model Report has not been developed for the combined Existing Facilities Program.

[7] NYSERDA, *New York Energy $mart^{SM} Enhanced Commercial Industrial Performance Program – Updated Program Logic Model Report*, June, 2007 and NYSERDA, *New York Energy $mart^{SM} Peak Load Management Program – Updated Program Logic Model Report*, July, 2007.

[8] Including utilities, the NYISO, and other third-party administrators.

NYISO DR program tracking databases, U.S. Census County Business Patterns Reports, membership lists and other publicly-available data from relevant professional organizations, and other sources identified and deemed valuable during a scan of relevant literature. Where possible, market characterization results will be segmented on an upstate-downstate regional basis to identify spatial variations in program and market opportunities and barriers throughout New York.

Example market characterization metrics to be developed pending data availability include[9]:

- EFP market share as measured by energy savings and peak demand reduction, segmented by utility territory and market sector, and eligible contractor base;
- Existing energy efficiency/DR service delivery channels (*i.e.*, energy service companies (ESCOs), curtailment service providers (CSPs), meter service providers (MSPs), load serving entities (LSEs), and/or other DR service providers) and most active contractors;
- Number of contractors active in the state, annual value of contractor activity, and market perceptions of industry strength and competitiveness;
- Current participation rates in NYISO and utility DR program offerings and Existing Facilities Program interaction with these programs;
- Current penetration of advanced metering infrastructure (AMI) and smart grid platforms; and
- Other metrics as identified.

Market Assessment

Market Assessment results will be generated through primary data collection efforts with end-use customers and contractors participating in the EFP as well as with comparison non-participant groups eligible to participate in the program (See next subsection for specific details regarding the proposed data collection efforts). The data collection instruments will be structured around the prioritized measurement indicators and researchable issues identified in the final *Program Logic Model Report*.[10] Care will be taken to ensure continuity of longitudinal indicator measurements where appropriate so that temporal trends in the measurements can be assessed. Market assessment results will be segmented on an upstate-downstate regional basis to identify spatial variations in responses and associated market conditions.

Example indicators to be measured during the market assessment work include:

- Market awareness of NYSERDA program offerings and broader energy efficiency and peak demand reduction opportunities;
- Market perceptions regarding the value of energy efficiency and DR services promoted by the program and the benefits associated with energy efficiency and peak demand reductions;
- Customer decision-making processes including organizational structure and financial and other non-energy considerations;

[9] Similar metrics and indicators will be used to assess the impact of program interventions on the broader contractor (*e.g.*, ESCO) market when the MCA Team turns its attention to the energy efficiency performance based incentive program components during the 2010 evaluation cycle. Evaluation results will enable the MCA Team to examine the extent to which program interventions are helping to develop a viable contractor industry in New York.

[10] Other evaluation contractors will be able to suggest additions to the instruments to collect data relevant to separate studies and the MCA Team will attempt to accommodate such requests, balancing the additional survey components against the need to minimize impacts on survey respondents.

- Market awareness of and participation in NYISO and utility DR program offerings;
- DR service provider expertise with peak demand reduction strategies promoted by the program including dynamic retail pricing and market-based pricing programs; and
- Other indicators as identified.

Analysis and Reporting

Data analysis and reporting will be conducted by the MCA Team using methods approved by NYSERDA. As discussed above, the analytic process will make use of both primary and secondary data sources to generate comprehensive and unbiased information regarding the market eligible to participate in the EFP as well as the success of program intervention strategies. All data sources used in the analysis and reporting phase of the project will be clearly cited to ensure a transparent record of activities undertaken. In addition, evaluation findings will be related back to the outputs and outcomes anticipated by the program logic model to help NYSERDA staff and other project stakeholders better assess program accomplishments to date.

Before preparing the final evaluation report, the MCA Team will present preliminary results to NYSERDA evaluation staff, EFP staff, and other project stakeholders to review key findings, clarify discussion points as necessary, and ensure accurate interpretation of results. Feedback generated during this presentation will be incorporated into the initial draft final report submitted to NYSERDA. An iterative process will then be used to finalize the report whereby the MCA Team will address feedback received during the report review cycle(s) until the report is deemed final by NYSERDA staff and other project stakeholders. Final evaluation results will also be presented to DPS and other project stakeholders during scheduled meetings.

Populations/Samples

As discussed previously, the MCA evaluation of the EFP is structured to accommodate the multiple components of the Program in that the MCA Team will initially focus its efforts on the Demand Response Performance Based Incentive Program component during the 2010 evaluation cycle, and subsequently turn its attention to the Electric and Gas Efficiency Performance Based Incentive Program components during the 2011 evaluation cycle. Each evaluation will involve primary data collection with end-use customers and contractors participating in the program as well as with comparison non-participant groups eligible to participate in the program.[11] The MCA Team will work closely with NYSERDA's data collection contractor to identify potential end-use customer and survey contractor sample frames and to develop sampling procedures to effectively represent the participant and non-participant populations.[12]

[11] The MCA Team will explore opportunities to aggregate primary data collection efforts across programs into sector-wide or market-wide efforts. Doing so may help 1) avoid duplication of effort in interviewing sets of market actors common to many programs (*e.g.*, ESCOs) and 2) hedge against the risk of overlooking certain market sectors not explicitly targeted by specific program offerings. In addition, the MCA Team will remain aware of the activities of the EAG's subcommittee on statewide studies to again avoid potential duplication of effort but also to determine how best to supplement any statewide studies approved by the DPS. Results of these efforts will be discussed in the final project workplan.

[12] The samples may be segmented by program component participation path (*e.g.*, segmented by Load Curtailment/Shifting, Distributed Generation, and Interval Meters for the Demand Response Performance Based Incentive Program component) or some other relevant variable (*e.g.*, anticipated peak demand reduction).

The final sample sizes for all market actor groups will be designed to meet 90/10 absolute confidence/precision criteria on an upstate-downstate regional basis.[13]

Current estimates regarding sample sizes, expected sampling precision, and anticipated survey fielding dates for the 2010 MCA evaluation are summarized in Table 3.[14] These estimates will be finalized prior to undertaking the planned evaluation, once the MCA Team more thoroughly analyzes program participation data, and will be documented in the final 2009 project workplan.

Table 3. Existing Facilities Program MCA 2009 Evaluation Specifics (Focused on the Demand Response Performance Based Incentive Program component)

Target Group	Estimated Population Size	Estimated Sample Size[1]	Expected Sampling Precision[2]	Survey Administration By	Expected Start of Fielding
Participating End-use Customers	TBD	140[a]	90/7	Survey Contractor	Winter-Spring 2010
Participating Contractors	TBD	140[a]	90/7	Survey Contractor	Winter-Spring 2010
Non-participating End-use Customers	TBD	140[a]	90/7	Survey Contractor	Winter-Spring 2010
Non-participating Contractors	TBD	140[a]	90/7	Survey Contractor	Winter-Spring 2010

[1] Final population and sample sizes will be determined through conversations with program staff and initial market characterization exercises. For budgeting purposes, it is assumed that the population size of each Target Group is large enough that 70 completed surveys will be required in each of the upstate and downstate regions (*i.e.*, 140 completed surveys per Target Group) to achieve 90/10 absolute confidence/precision criteria on an upstate-downstate regional basis.

[2] Assumes proportional sampling, two-tailed test, finite population correction

[a] Should NYSERDA be directed that data collection efforts achieve 90/10 confidence/precision levels on a utility territory basis, the sample sizes and associated data collection costs will increase accordingly. If this occurs, the results would benefit all EEPS program administrators and NYSERDA would propose that the data collection efforts be undertaken in a jointly-funded manner with all program administrators contributing.

Data Collection

Primary data collection with each market actor group will be managed by NYSERDA's survey contractor. The data collection process will be conducted by telephone[15] and will consist of the following

[13] Should NYSERDA be directed that data collection efforts achieve 90/10 confidence/precision levels on a utility territory basis, the sample sizes and associated data collection costs will increase accordingly. If this occurs, the results would benefit all EEPS program administrators and NYSERDA would propose that the data collection efforts be undertaken in a jointly-funded manner with all program administrators contributing.

[14] Similar estimates were used to develop budget estimates for the proposed 2011 MCA evaluation. Final metrics, including corresponding budget estimates, will be developed prior to launching the 2011 evaluation and will be documented in the final 2011 project workplan.

[15] Surveys will be designed to be completed in approximately 15 to 20 minutes.

steps undertaken by the survey contractor: 1) Format the final survey instruments and program them into a Computer Aided Telephone Interviewing (CATI) system, 2) Pretest the final instruments with subsets of the market actor group samples and consult with the MCA Team as needed to resolve any issues that are identified[16], 3) Conduct full-scale data collection efforts and provide regular progress updates to the MCA Team during implementation, 4) Process the raw survey data into final data files including coding of open-ended responses and general data cleansing, and 5) Deliver to the MCA Team final data files in SPSS and SAS formats including all variable names, variable labels, value labels, and weights relevant to each data collection effort along with the associated codebooks.

The MCA Team will coordinate with NYSERDA's other evaluation contractors to the extent possible to fully leverage other planned data collection efforts. Such coordination efforts will achieve economies of scale in terms of minimizing data collection costs, ensure consistency of approach and question wording to facilitate comparison of results across evaluation efforts, and minimize the burden placed on different respondent groups. In addition, the MCA Team will work closely with the impact evaluation contractor team to ensure that final MCA results are considered during the attribution analyses conducted by that team (see discussion in Section 6). The EFP is designed to have a strong market transformational aspect, and the theory-driven results generated by the MCA evaluations will ensure the program is credited for structural and functional changes in the market that result from program interventions, changes that market actors contacted during attribution analyses may not be fully cognizant of.

The proposed MCA evaluation schedule and budget for the EFP are shown in Table 4. These initial budget estimates will be finalized prior to undertaking the planned evaluations and once the MCA Team more thoroughly analyzes program participation data.

Table 4. **Existing Facilities Program MCA Evaluation Schedule and Budget**

Evaluation Element	Estimated Budget and Completion					
	2009	2010	2011	2012	2013	Total
Market Characterization & Assessment	--	$203,000a	$184,000a	--	--	$387,000

a Primary data collection costs represent approximately 37% of the total proposed evaluation budgets.

VI. Impact Evaluation Plan

The Enhanced Commercial and Industrial Performance Program (ECIPP) and Peak Load Management Program (PLMP) were combined to create NYSERDA's Existing Facilities Program (EFP). Together, these programs were anticipated to complete approximately 340 projects per year during SBC III. While only 16% of SBC III projects were located in the downstate region as of June, 2009, these projects are expected to provide about 44% of the energy savings. The 2010 evaluation will examine ECIPP and PLMP SBC-funded projects completed in 2006 through 2009. (Gas measure impact is the subject of a separate evaluation.)

[16] Pretest interviews will be included as completed interviews unless major revisions to the instruments are made.

NYSERDA began a targeted initiative towards industrial process efficiency projects in mid-2008 as part of the preceding ECIPP. This effort has been incorporated into a new module within Existing Facilities, the Industrial Process and Efficiency Program. This program is being evaluated separately from the EFP, and is not part of this evaluation.

Research Objectives

The purpose of impact evaluation is to establish rigorous and defensible estimates of the savings that can be attributed to the efficiency program. One part of this process is to determine the realization rate, *i.e.*, the ratio of the actual verified gross savings to the NYSERDA-reported gross savings (*ex ante* savings estimates). The net effects of the program (attribution) are also necessary to separate the program impacts from naturally occurring efficiency. In both of these aspects of the impact evaluation, the evaluators need to determine how to achieve the desired precision, minimize the possibility of bias in the result, and assess the validity of the results. Each of these key aspects of impact evaluation is discussed briefly below.

Determine Realization Rates for Gross Savings

A critical component of the impact evaluation is to develop rigorous estimates of the realization rates for gross electricity, demand, and fuel savings, which will entail verifying the installation and the estimation of savings for a representative sample of program participants in comparison to an established baseline. The gross savings realization rate is then applied to the program population of NYSERDA-reported gross savings to derive the final savings estimates (evaluation-based estimates, or *ex post* savings).

Attribution

An equally important element of assessing impacts is to construct solid and defensible estimates of all impacts that are program-induced (rather than naturally occurring). This is often accomplished through estimation of the ratio of impacts for those that would have taken the actions without the program (free-riders) compared to program savings and the ratio of the savings from actions taken outside NYSERDA programs but due to the program (spillover). The combination of these components in the form of a net-to-gross (NTG) ratio becomes the adjustment factor to derive net savings.

For EFP, net effects will be measured for participating customers and vendors. This assessment of net effects will cover participant spillover. Non-participant spillover could easily overlap with NYSERDA's other programs targeting the C/I sector (*e.g.*, Flex Technical Assistance, Business Partners), and these effects are planned to be measured through a study of the entire C/I existing facilities market scheduled for 2009. Future estimates of non-participant spillover could be accomplished as joint studies with all EEPS program administrators to ensure consistent methods and avoid overlap in claimed savings.

Precision and Bias

Sample sizes will be designed to target 90/10 precision or better for the combined Existing Facilities program (PLMP and ECIPP), for the upstate and downstate regions.[17] The primary evaluation outcomes

[17] For the purposes of this document we classify all projects as either PLMP or ECIPP. PLMP is intended to encompass all demand reduction projects as well as EF projects that would have been considered PLMP projects the programs had not merged to form EF. "ECIPP" encompasses all other efficiency projects, including CHP.

will be specifically identified for these precision targets as the sample design is specified.[18] The proposed analysis methods will also be considered as part of developing the sample size requirements. Methods will be selected to minimize self-selection, non-response, and other sources of bias, to the extent possible.

Activities

Gross Savings Impact Evaluation

The EFP is a core NYSERDA program accounting for more than 10 percent of the SBC portfolio funding and savings. It is the compilation of two prior NYSERDA programs, Enhanced Commercial Industrial Performance Program (ECIPP) and Peak Load Management Program (PLMP)[19] that were active in the 2006 to 2008 period, plus an earlier program, Smart Equipment Choices (SEC). Existing Facilities has diverse program elements, as described below:

- Pre-qualified efficiency incentives, gas & electric (ECIPP Tier I and PLMP Permanent Demand Reduction (PDR) Small Lighting)
- Custom efficiency incentives, gas & electric (ECIPP Tier II). This option enables customers to receive incentives on smaller projects that otherwise would need to be performance-based by using pre-approved savings calculators
- Performance-based efficiency incentives for electricity, gas, and steam (ECIPP Tier III and PLMP PDR):
    - i. Commercial
    - ii. Industrial process (not in the scope of this evaluation)
    - iii. Building management systems
    - iv. Steam-driven chillers, for Con Edison steam customers
- Demand reduction incentives to enable participation in the NYISO DR programs, including:
    - v. Interval meters
    - vi. Distributed generation for participants with existing emergency or back-up generators
    - vii. Load curtailment or shifting
- Combined heat & power incentives

The performance-based CHP and DR elements, with the exception of interval meters, require pre-project metering to demonstrate savings potential. Most ECIPP Tier III and CHP projects also require post-installation measurement and verification (M&V) to demonstrate savings. NYSERDA's implementation team employs third-party technical assistance (TA) contractors to review the implementing contractors' proposed M&V plans and subsequent reports as part of program quality assurance (QA).

---

[18] The sample size depends on the type of statistical analysis being conducted and the variability of the specific parameters to be estimated. For example, a simple random sample required to achieve a 90% confidence and 10% sampling precision for a yes/no question is about 67 for a large population. However, if the variable of interest is the realization rate and the coefficient of variation is 0.75, a simple random sample would require a sample size of 152 to achieve the same precision and confidence level.

[19] This Peak Load Reduction program became PLMP in SBC III as its objectives expanded to include peak load management, operations and behavioral actions.

In general, prior evaluations have found that projected savings for the legacy **New York Energy $mart**[SM] Programs that are the basis for this program use sound engineering calculations and rigorous post-installation verification activities. At the same time, the increased evaluation funding can substantially add to the overall reliability in the independent evaluation savings estimates by funding significant expansions in the M&V methods. More sophisticated methods with greater measurement support can greatly reduce any unknown risks of potential bias that can go unobserved within more simplistic methods.

2010 Impact Evaluation

The 2010 evaluation will examine gross impact and attribution for program activity during 2006 through 2009. The evaluation will be divided into five components:

- permanent demand reduction for PMLP participants (M&V site visits for a sample)
- interval meters for PLMP participants (analysis of metering records on a sample)
- steam retention, load curtailment, DEGI for PLMP (M&V site visits on a census)
- PLMP persistence study for permanent load reduction measures
- ECIPP (M&V site visits for a sample, possible billing analysis on a subset)

The basis for this approach is described in more detail below.

PLMP has five distinct components that will require different evaluation strategies. Table 5 shows the activity in each component from 2006 through June of 2008. This analysis shows that the interval meters (IM) and permanent load reduction efficiency (PDRE) are by far the most heavily used components of the PLM program. Further review of the program data indicates that 53% of the kW reduction associated with the PLMP during 2007 and the first half of 2008 was due to the IM component, while the PDRE component accounted for 27% of the kW reduction.

Table 5. PLMP Participation for SBC-Funded and Partially SBC-Funded Projects[1]

PLMP Component	Number of Participants By Year			
	2006	2007	Through Q2 2008	TOTAL
Permanent Demand Reduction (PDRE) – Efficiency	74	406	142	**622**
PDRE – Steam Retention	6	1	0	**7**
Load Curtailment/Shifting	4	4	1	**9**
Dispatchable Emergency Generation (DEGI)	2	10	1	**13**
Interval Meters (IM)	15	287	5	**307**
Total	**101**	**708**	**149**	**958**

1. A portion of these projects also received funding through the System Wide Program.

Persistence of demand reduction projects is also a potential issue with the PLMP program. Participating contractors are generally required to deliver demand reduction for at least two years after project completion and to self-report on delivery. All project funding is paid as soon as installation is verified. The penalty for failure to deliver the requisite reports or at least 80% of committed demand reduction is exclusion from future projects; no funds are contingent on sustained performance.

ECIPP as implemented from 2006 through 2009 has four distinct components: Tier 1 pre-qualified measures, Tier 2 custom measures with savings based on engineering methods, Tier 3 custom measures with savings developed from performance-based measurements, Combined Heat and Power and Smart Equipment Choices. These projects may be combined for sampling and analysis purposes.

Impact Evaluation Approach

The Impact Evaluation Team will consider a variety of approaches for evaluating the Existing Facilities Program including the following:

- Billing analysis;
- IPMVP Option A: M&V spot measurement and verification; and IPMVP Option B: retrofit isolation measurement with data from some or all of these sources:
- Evaluator direct short-term measurement;
- Leveraged use of participant energy management system data; and
- Leveraged use of data collected by the NYSERDA-funded third party M&V contractor.

Impact evaluation is likely to include billing analysis for more homogeneous groups and significant site survey work and engineering analysis with calibration to billing records for a sample of the more heterogeneous projects (*e.g.*, IPMVP Option B: Retrofit Isolation where complete measurement is possible). Alternatively, independent site M&V will support the use of IPMVP Option A by undertaking metering/monitoring measurement to mitigate the greatest sources of uncertainty, as appropriate. The projects that participate in the new Building Management System module, proposed in the EEPS 90 Day

plan under EFP, provide important post-retrofit monitored data that can be leveraged for evaluation.[20] The program is also funding significant third-party baseline plus post-retrofit measurement and verification that evaluators plan to leverage.

The advantage of a billing analysis is that it allows for larger sample sizes or analysis of census of projects within a particular group, improving the rigor and precision of the results. It also allows for the possibility that savings could be calculated at the utility service territory level with 90/10 precision or better and it automatically accounts for snapback. However, billing analysis is not an effective method of estimating savings for highly diverse projects. In this case, individualized M&V plans will need to be developed, and billing records will instead be used as a key input for calibrating savings to actual use.

The first step in this preparation for the impact evaluation will be to review the program data to determine the range of projects and assess which projects may be included in a billing analysis, compared to those that are more appropriate for an individualized M&V approach. For example, if many lighting measures were installed in small- or medium-sized offices, it may be appropriate to develop a statistical model to estimate average lighting savings for these establishments. The feasibility of a billing analysis is predicated on the magnitude of the savings (typically 8% to 10% of annual consumption, although smaller effects can be found with larger sample sizes or highly homogenous groups), the availability of billing records for the entire analysis period starting one year prior to the installation and extending to a full year following it, and access to information on economic trends that may affect energy use.

Given the changing economic climate, it may be advisable to combine the billing analysis with a telephone survey to identify other key factors that affect energy use, such as changes in the number of employees or in the hours of operation. This approach reduces the potential sample size for the billing analysis since it restricts the sample to those participants who complete the telephone survey rather than allowing the inclusion of all participants with sufficient billing records, but it also provides critical supporting information to interpret the billing data. Fielding the telephone survey allows for a comparison of techniques in that the billing analysis can be conducted using all participants with sufficient billing records (without the additional telephone survey data) and only for the sample of participants who responded to the telephone survey.[21]

A sample of the larger savings and more diverse projects will be selected for individualized treatment. Sampled EFP participants will be subject to one of three levels of evaluation to determine the program gross savings realization rate:

- Inspection or review-only verification, for the smallest savers;
- IPMVP Option A-level analysis if the project delivers moderate savings and the evaluation engineer finds that the implementation-side M&V was conducted in a sufficiently rigorous and objective manner to permit leveraging the data; and
- IPMVP Option B or D-level analysis, for all large savings projects in the sample and for moderate savings projects that lacked prior evaluation-grade analysis through the program.

[20] However, if should be noted that this evaluation plan is only intended to cover the current EFP, and may need to be revised/expanded should NYSERDA receive approval of the EFP per the EEPS 90-day filing.

[21] The proposed evaluation plan and budget currently do not include the telephone survey option for a subset of participants for this alternative billing analysis.

To the extent possible, the gross savings findings related to the pre-qualified measures may also be incorporated into NYSERDA's Deemed Savings Database.

Projects with performance-based incentives are required to have site-specific M&V plans under the current program design. In some cases, these M&V plans will incorporate independent third-party measurement, funded by NYSERDA and paid directly to the Technical Assessment contractor. The Impact Evaluation Team will work with NYSERDA and these contractors to incorporate evaluation-grade techniques into the scope of these task work orders. The 2010 evaluation effort will include exploring the extent to which this effort can be drawn upon for the evaluation to maximize evaluation cost-efficiency and customer service for later evaluation cycles.

When possible, the evaluators will avoid replicating the review and installation verification activities already completed by NYSERDA's independent third party review contractors. In most cases, however, such activities have not included metering which will be necessary to obtain rigorous savings estimates. The strategies to be applied to specific end uses are described below.

- Lighting. A material number of developers have participated with prescriptive lighting projects. Fixture counts have already been verified by independent third parties. Such projects are expected to be subject to lighting logger studies but not power metering (unless lighting panel instrumentation is more cost-effective) or re-verification at large sites.

- Envelope. Eligible envelope measures such as heat island mitigation typically will be modeled with building simulation tools.

- Renewable generation. Analysis of renewable generation-based PDR will rely on participant data when data availability and quality allows.

- Motors, chillers, drives, and related measures. M&V of savings due to high efficiency chillers, motors, and similar measures typically will follow IPMVP Option B protocols.

The analysis will follow standard M&V practices and will vary according to project type. The following paragraphs include a brief description of each of the five components of the evaluation.

PLMP Permanent Load Reduction

Individual M&V plans will be developed for each site depending on the specifics of the project. The recent Large Savers evaluation indicates that the engineering estimates of savings for the PLMP projects tend to be quite accurate, at least for the large projects included in that study. The Impact Evaluation Team will leverage direct measurement data from NYSERDA's contractors as available and appropriate.

The PLMP evaluation will include an analysis of non-energy benefits. NYSERDA began estimating values for participant non-energy impacts (NEIs) in the 2003-2004 evaluation cycle. NEIs have not previously been estimated for the PLMP, nor are they included in the benefit-cost analysis. There may be significant and directly quantifiable NEIs for PLMP participants, such as more favorable tariffs for steam retention and load curtailment participants and higher or lower Operations and Maintenance costs. NEIs that are more difficult to quantify, such as those traditionally associated with efficiency projects (marketing benefits, employee retention, etc.), also apply to some PLMP projects. Deriving NEI estimates for this program will put it on equal footing with NYSERDA's other major commercial and industrial programs. NYSERDA proposes to perform direct query-based NEI evaluation and possibly conjoint method-based evaluation on SBC-funded PLMP projects as part of the 2010 impact evaluation

on programs years 2007 - 2009. The sample for NEI evaluation will be the same as or a subset of the gross savings evaluation sample.

Evaluators previously analyzed NEIs for the former ECIPP in 2006 using both the direct query and conjoint methods. Thus, the evaluation of NEIs will focus on the PLMP in this evaluation cycle.

Steam, Load Curtailment and Dispatchable Generation (PLMP)

Given the small number of projects in the steam, load curtailment and dispatchable generation components of the program, the Impact Evaluation Team intends to conduct a census attempt for these participants. Individual M&V plans will be developed for each site depending on the specifics of the project. With the exception of the steam project, the Impact Evaluation Team expects that the PLMP site-specific plans will not require an intensive engineering effort. For example, the DEGI project may simply require establishing the operational status of the back-up generator, the available capacity of the generator and a review of billing records to ascertain the magnitude of the potential curtailable load.

Interval Meters

The installation of the interval meters allows customers to participate in the NYISO and utility demand reduction programs. Previous evaluations indicate that the demand reduction was estimated based on self-reports from the participants. Program Opportunity Notice (PON) 1097 requires that each IM contractor commit to providing NYSERDA access to interval meter data for a minimum of two years after installation. In 2010, evaluators will request that the project developers provide this data, or access to it, for gross demand and energy savings analysis of sampled sites. The energy savings analysis will include examination of post-control period snapback. Most IM projects are subject to verification by NYSERDA's independent third party review consultant. If the review consultant did not perform on-site verification of meter installation (*e.g*. it was a one- or two-meter project) and data analysis indicates a need for it, the evaluator will verify the installation. However, the Impact Evaluation Team shall set a minimum threshold for performing a site visit according to the amount of NYSERDA incentives awarded.

The savings associated with interval meters is based on the ability of the participant to reduce demand in response to a request from the NYISO or the utility. Since these requests are likely to be driven by high coincident peak demand, it is likely that these events will take place during periods of high activity. Consequently, the savings need to take into account the actual potential demand reduction given the level of activity that must be maintained at the site. If a site has yet to be requested to respond to a call for demand reduction, it will be excluded from the sample.

Persistence

If participants are able to deliver demand reduction savings for the first two years after receiving funding, this does not necessarily indicate that they will still be prepared to respond to an NYISO or utility event three or four years after participation. A component of this evaluation will be to select a stratified random sample of older projects with curtailable load reductions to assess whether the load management procedures are still in place and operational. The availability of the load curtailment will be evaluated through a telephone interview, followed by a site visit at up to 10% of the sites.

ECIPP

For ECIPP contractors, the scope of work typically covers independent review of the project developer's submitted M&V plans and calculations, checks of the claims for reasonableness, and in some cases

verification of the installation of equipment. Past impact evaluations have relied heavily on these reviews due to budget necessity. In contrast, a more rigorous impact evaluation depends on third-party M&V of savings based on independent interpretation of data and site measurement as needed. As with PLMP PDR projects, individual M&V plans will be developed for each site depending on the specifics of the project. The Impact Evaluation Team expects that the site-specific plans for ECIPP custom projects will not require an intensive engineering effort since direct measurements should be available from NYSERDA's contractors. The Impact Evaluation Team will review the program data to assess whether any subset of participants with similar conditions and measures could be combined into a billing analysis.

Attribution

The savings attributed to an efficiency program should be the savings induced by the program effort, above and beyond what would have occurred in the absence of the program. For program participants, assessing attribution involves estimating the program measures (or the proportion of the savings) they would have adopted within the same time frame but without the program (free-ridership. Program participants can also take additional efficiency actions due to what they learned or experienced through the program even when these actions were not recommended in the program report (spillover). There are two types of participant spillover:

- “Inside” spillover occurs when additional actions are taken to reduce energy use at the same project site, but these actions had not been program measures (included within the incentives provided); and
- "Outside" spillover occurs when an actor participating in the program initiates additional actions that reduce energy use at other sites but due to the NYSERDA program.

In addition, the program can also influence non-participants. The influence of NYSERDA’s EFP on the commercial, industrial and institutional sectors can easily overlap with the influence of the NYSERDA’s other commercial and industrial programs. Recognizing this, NYSERDA conducted a commercial and industrial (C/I) non-participant spillover study applicable across most of its C/I programs in 2005 and updated in 2007. NYSERDA plans to conduct a similar but expanded study in 2009 to derive updated non-participant spillover rates for all its C/I programs. Consequently, this component of the Existing Facilities NTG effects will not be addressed in this EFP evaluation.

The Impact Evaluation Team intends to explore participant free-ridership and spillover for the EFP with participating customers and the vendor or ESCOs that solicit and manage the projects.[22] Other decision-makers, such as chief financial officers or vendors, may also be interviewed for the largest projects in both program areas, if they are found to be highly influential in the decisions to invest in energy efficiency measures. The Impact Evaluation Team plans to leverage the interviews required for the gross savings evaluation for the NTG study and minimize the overall contacts with the participants.

To the extent possible, questionnaire design and triangulation methods will test and address the potential bias that may occur from interviewing participating contractors where self-interest would suggest that responding in a manner to provide a high NTGR is more likely to lead to continuation or expansion of the program. If a customer states their reason for assessing and adopting a measure was due to the influence of the participating contractor, the contractor might have offered and promoted the measure without the program. Alternatively, the contractor might not have existed without the market development that the program encouraged to occur. An opportunity to better understand these complexities and the inter-

[22] Customers can also propose projects to NYSERDA and lead the projects themselves. There are also pre-qualified measures that have simpler application and lower project measurement and verification (M&V) requirements.

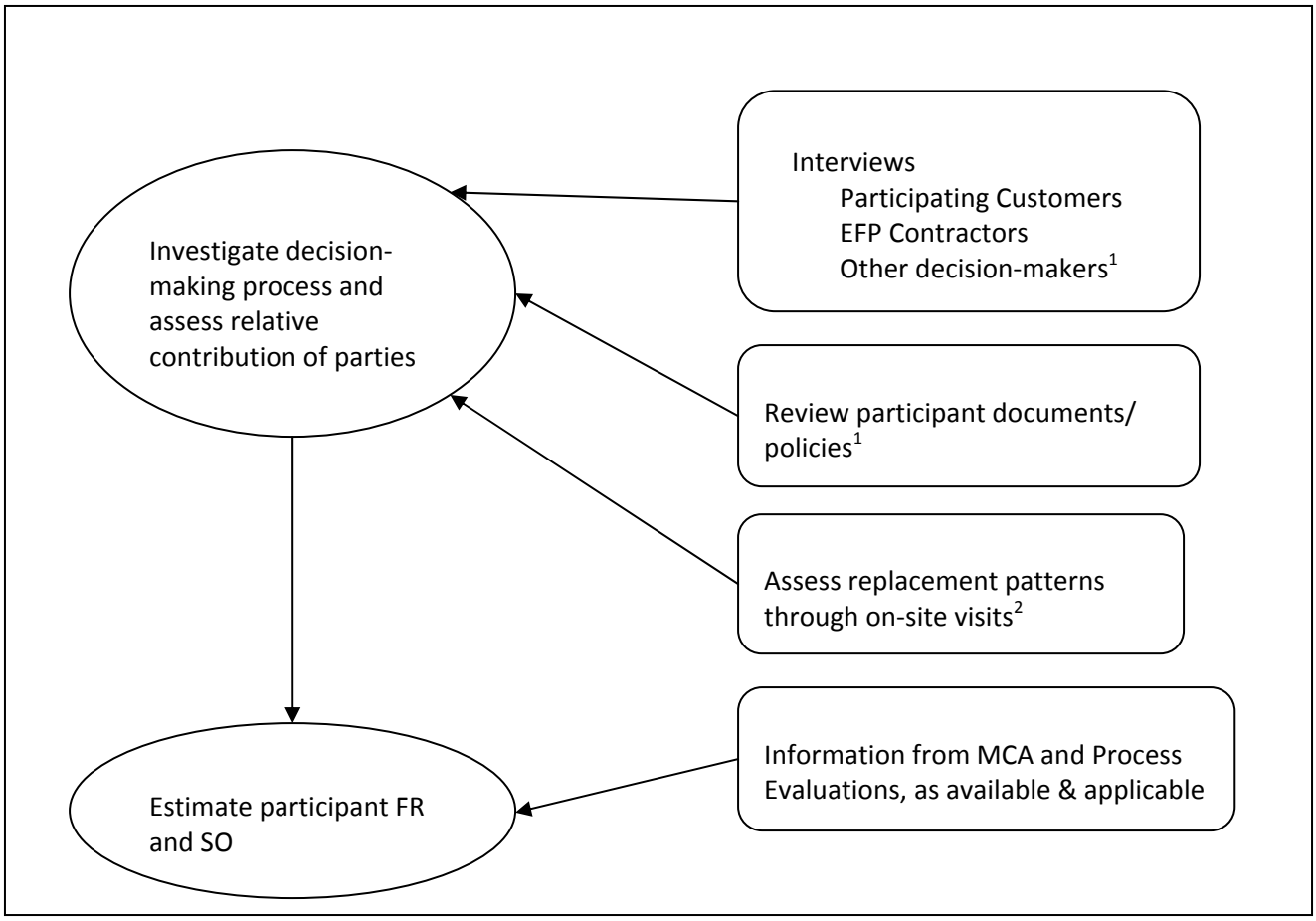
relationships between the adoption decisions and market development will be explored with the MCA evaluation team, working with the design of their market study that is also being planned for 2010.

The discussion of sample sizes is included below in the section on population/samples. The reliability for attribution, however, relies more on construct validity than on sampling precision. The alternative of what would have occurred cannot be known with certainty. Survey inquiry can be difficult when asking about conjecture of a theoretical alternative. Prior survey experience for specific question wording, measuring free-ridership in more than one way, and obtaining market or other comparatives are several ways to increase the reliability of the attribution estimate. The market studies may provide useful comparatives. Similarly, interviews with the largest projects could be supplemented with internet research on that organization's criteria for similar investments. Measuring free-ridership in multiple ways can increase the construct validity of the estimate.

If the statewide baseline studies being proposed by NYSERDA for consideration by all EEPS program administrators are undertaken, these results will be analyzed and compared to the Existing Facilities participant sample to derive a market-based estimate of naturally-occurring adoption as an alternative NTG input value. Technologies that are part of any available recent cross-state comparison market effect studies will be compared with the combined results from the TA program and NYSERDA's other C/I program evaluation results. These alternative methods will be combined with the participant enhanced self-reports described above to derive a draft triangulated NTG ratio. These draft NTGR results will be reviewed and discussed, along with the Impact Evaluation Team's recommended triangulation method, with DPS staff and the NYSERDA evaluation project manager. Based upon comments received in this review, the Impact Evaluation Team will finalize the EFP free-ridership and participant spillover estimates. Using multiple methods and data sources will provide a high level of construct validity for the NTG estimates.

Figure 4 illustrates the sources of information that the impact team will use to analyze attribution.

Figure 4. Participant Free-Ridership and Spillover



1. To be conducted on census strata participants in the sample.
2. For those participants receiving site visits as part of the gross savings evaluation.

Survey design and solicitation will be designed to minimize self-selection to the extent possible.

Obtaining cooperation and good response rates from customers for the NTG surveys can be difficult with most commercial and industrial customers, especially firms that undertook custom projects and have already been subject to extensive program M&V and review through a quality assurance (QA) process. The Impact Evaluation Team and NYSERDA staff will explore alternative methods to address this challenge in 2010, with some preliminary improvements to be tested in the 2010 evaluation.

DPS will have the opportunity to review all NTG survey instruments prior to field use.

Populations/Samples

Sampling is necessary to estimate both gross and net impacts, as discussed in more detail below. Since the legacy programs were implemented separately during the 2006-2008 program years separate samples will be drawn to meet those needs.

<u>Gross Impact Sampling</u>

The sample will be designed to determine the realizations rates for the upstate and downstate regions separately at the 90/10 confidence/precision level. It is not designed to provide separate realization rates

for the PLMP and ECIPP components of the program at the 90/10 level. While realization rates for the separate components may be calculated and reported, if desired, the relative precision is unlikely to meet the 10% target and may be substantially worse. Given the high costs of conducting size visits for these types of projects, this approach provides the balance between maintaining the desired precision level and keeping evaluation costs within a reasonable range.

Efficient sample sizes will be chosen using stratified ratio estimation (SRE). Two samples will be drawn for the gross impact evaluation, one for the upstate regions and a one for downstate. Projects completed in 2006 through 2009 will be included in the sample frame. The Impact Evaluation Team will attempt to reach a census of the PLMP steam, load curtailment and DEGI projects.

For the upstate region, 16 PDR and 60 interval meter projects will be randomly selected for PLRP as well as 43 projects for CIPP. The downstate sample will include 16 PDR and 40 IM projects in the PLMP and 16 CIPP projects, in addition to the 20 PLMP steam, load curtailment and DEGI projects in the census stratum. The sample sizes were based on historical savings data (kW) for projects completed in 2006 through mid-2008, assuming an error ratio of 0.6 and the previously-stated target confidence/precision levels for the upstate/downstate regions.[23] The precision targets were adjusted by component within the sample to obtain the target value for the upstate and downstate samples, taking into account the percent of savings associated with each component and the absence of sampling error for the census component. The sample sizes for each component are shown in Table 6 below.

[23] Among PLMP PDRE projects, previous evaluation activities suggest that the error ratio is about 0.30. The error ratio of 0.5 was intended to allow for differences in evaluation strategies and provide some leeway in the sample sizes.

Table 6. Sample Sizes by Program Component

	Totals	IM	PLRP PDR	PLRP Census	ECIPP
Upstate					
PLMP	16		16		
PLMP IM	60	60			
CIPP	43				43
Totals	119	60	16	0[a]	43
Downstate					
PLMP	16		16		
PLMP IM	40	40			
PLMP Census	20			20	
CIPP	17				17
Totals	93	40	16	20	17
Grand Totals	212	100	32	20	60

a There were no PLMP census projects in the upstate region.

In addition, a third sample of 70 older curtailable load PLMP projects will be selected to assess persistence. This study will be conducted for the PLMP only and will not be designed to produce separate persistence rates for activity in the upstate and downstate regions. An SRE sample will be drawn from the load curtailment projects completed during 2004 or earlier. Previous evaluations do not provide a basis for estimating the sample size. The Impact Evaluation Team assumes that the error ratio will be in the range of 0.50 and a sample of 70 projects will be sufficient to assess the impacts of persistence at the 90/10 level. This approach will not be designed to assess whether the realization rates change from one year to the next, but rather to determine the realization rate for the older (pre-2005) projects for the program components designed to achieve curtailable load reductions.

All samples will be stratified by size (kW savings). A census of the high energy-saving sites in the top stratum will be reviewed and a random sample of sites within each of the lower strata will be selected for the sample. The smallest savers may be eliminated as site visit candidates. Results will be weighted by stratum and for the program as a whole.

Attribution Sampling

The evaluation of net impacts is focused on participating customers, vendors and ESCOs. For efficiency and to avoid survey fatigue, the 2010 gross and net surveys may be integrated for participating customers. In addition, the Impact Evaluation Team will consider whether the engineering firms, ESCOs and vendors associated with the selected projects will constitute a sufficient sample. When both the participating customers and ESCOs for the same projects are surveyed, it allows the Impact Evaluation Team to gain valuable insights into how the same project is viewed from these multiple perspectives. Particularly when the Impact Team has been on site and has first-hand knowledge of the actual installations, this approach provides useful validation of the self-reported net effects. However, the possibility exists that the sample

of site owners may not result in a representative sample of the more commonly-used and larger ESCOs or IM vendors. The final decision will be determined after a careful assessment of the composition of the sample, likely complexity of the decision-making process and extent of contact already occurring with these customers. For planning purposes, the Impact Evaluation Team assumed that the sample of 32 PDRE projects, 100 IM projects, and 60 ECIPP projects selected for the gross savings estimation will result in telephone surveys of 30 PDRE ESCOs and all 25 IM vendors (as estimated from previous evaluations).

The ECIPP attribution study also will be integrated with gross field work. A stratified random sample of 70 ECIPP projects will be selected to meet the 90/10 confidence/precision level for the program. Because the attribution sample size of 70 exceeds the gross savings evaluation size of 60, evaluators will need to collect attribution data for 10 projects independently of a gross savings evaluation. The sample frame will be stratified by the magnitude of savings. A census attempt will be made for the projects in the stratum with the highest savings, and a random sample of sites in the lower strata will be selected for the sample. The census strata for this study will use a more in-depth approach to estimating NTG factors than the lower strata and the site owners, site decision-makers and ESCO/vendor/engineering firms associated with the largest projects will be interviewed by social science professionals. For the projects in the lower strata, the interviews will be conducted by the survey contractor.

Review of the prior ECIPP market characterization, market assessment and causality evaluation reports indicates that there were about 170 ESCOs that participated in this program in 2005. Consequently, a sample of 70 ECIPP projects seems likely to capture many of the major players. However, if it appears that the ESCOs associated with these projects are not representative of the population of ESCOs/vendors/engineering firms participating in the EFP, the Impact Evaluation Team will consider drawing a separate sample from this population.

Data Collection

NYSERDA's impact evaluators will collect program tracking data from the implementation team, conduct independent surveys, and install instrumentation to collect independent data. The Impact Evaluation Team also plans to make extensive use of additional channels and resources not routinely used in prior NYSERDA evaluations, such as utility-provided bill data and continuous Web-based system performance required of certain program participants. This section details the impact evaluation data collection plans.

<u>Gross Savings Data Collection</u>

Approach. Engineers will perform field instrumentation to determine the gross savings realization rate, defining the instrumentation requirements in M&V plans in accordance with IPMVP terminology. They will estimate and consider engineering uncertainty and the cost associated with increasing or decreasing it for each plan. The approaches to be used are listed below:

- Short-term equipment metering of all key parameters, extrapolated to annual performance, with or without pre-retrofit metering (IPMVP Option B);
- Use of program-collected long-term energy performance data (IPMVP Option B);
- Spot and short-term metering of selected parameters (IPMVP Option A);
- Building or system energy use simulation, calibrated to bill data, equipment submetering or both (IPMVP Option D);

- Site-specific billing analysis (IPMVP Option C); and
- Prior savings claims methodology and data review (Engineering review).

The site plan may not strictly fall within the descriptions of the described approaches if pre-reviews indicate that available program data is objective and can be leveraged to maximize evaluation cost-effectiveness. Participant and developer interviews will supplement observed data. The data collection approach will be modified as necessary if a statewide protocol is established, as is anticipated in the *Evaluation Plan Guidance for Program Administrators* (Evaluation Plan Guidance).[24]

It is not possible to instrument all sampled sites simultaneously and during the summer peak capacity period. Engineers will attempt to have the measures with the most weather or seasonally-sensitive behavior instrumented during peak periods and measure other equipment at other times.

The evaluation will investigate potential overlap between this NYSERDA program and the New York ISO. This evaluation will ask participants about their participation in both programs and will also attempt to work with the ISO administrators to perform a database matching exercise to identify dual participants. All registrants in the NYISO demand response programs must have metered data available in order to participate. Customers to the PLMP program may apply for the interval meter under one project, then for the actual demand response equipment under another. Program implementation involves a step to check for past projects at the same site. When the match is found, the savings for the interval meters measure is reset to 0 kW. A second check is performed routinely to search for matches that may have been missed. The Impact Assessment team will use the database information on multiple measures at the same site to inform their site visits.

Resources. Evaluators will collect much of the data directly through measurement, interviews and review of existing program tracking data. To meet the level of rigor described in the Evaluation Plan Guidance, evaluators also will need utilities to provide billing data for participants and potentially a sample of non-participants. In addition, evaluators will request savings estimates developed by implementation-funded third party technical assistance contractors, applicants and their vendors.

To be able to conduct the sampling and proceed with the evaluation, the Impact Evaluation Team will need the following information at a minimum, in addition to primary data collected:

- Project level information, including address, contact information for the site owner, architect and engineer, the type of project (custom, design/build), type of business;
- Measure level information, such as a description of the measure, validated quantity installed, the energy savings (electric, gas and other fuels), demand savings, measure life, and incremental costs;
- Utility consumption data (electricity and natural gas), covering the date of the read, account number, premise number, amount of energy used, tariff, rate class, whether the read was estimated or actual, city or zip code, and weather station;
- Weather data, which may be available from the utilities or from the National Weather Service. If not available it can be purchased at an additional cost; and

[24] DPS staff issued these guidelines for incorporation by NYSERDA, electric and gas utilities, and third party administrators into their Electric and Gas Energy Efficiency Program proposals pursuant to the Commission's June 23, 2008 Order Establishing EEPS in Case 07-M-0548. The Guidelines were formed with input from the Evaluation Advisory Group and serve as an important first step in establishing evaluation protocols for the State's energy efficiency programs. They are available for download at http://www3.dps.state.ny.us.

- Excel workbooks, building and system simulation input files (.PD2 and .INP files in the case of DOE2), and other documentation of savings calculations developed by applicants and TA review contractors.

The population is defined as those projects that were completed during the 2006-2009 program period. As such the sample is post-retrofit only by design. For any project that is not new construction or an analog, pre-retrofit performance estimation is necessary to establish baseline conditions. Evaluators will collect data to estimate baseline energy use by either extracting the data from the customer's historical energy management system logs or by using the original data used in the application. If the latter, the evaluator will perform independent savings analysis. In some cases estimates will need to be made based on interviews.

In addition, critical information will need to be collected from third party sources, as detailed in Table 7.

Table 7. Existing Facilities Impact Evaluation Survey Specifics

Target Group	Estimated Population Size (2007/2008)	Estimated Sample Size	Expected Sampling Precision	Survey Administration By	Expected Start of Fielding
Participating IM Customers – PLMP- Interval data analysis, 10% with site visits Gross and NTG	~300	75	90/10	Impact Evaluation Team	Winter 2009-2010
Participating PDRE Customers - Site Reviews/On site Survey Gross and NTG	~600	32	90/10	Impact Evaluation Team	Summer 2010
Participating PDR Steam, LCS, and DG Customers - Site Reviews/On site Survey Gross and NTG	~25	20	Census attempt	Impact Evaluation Team	Summer 2010
Participating PDRE Customers - Site Reviews/On site Survey Persistence	~300	70	90/10	Impact Evaluation Team	Spring-Summer 2010
Participating PDRE Vendors/ESCOs – PLMP	70	30	90/10	Survey Contractor	Spring-Summer 2010
Participating IM Vendors – PLMP	25	25	Census	Survey Contractor	Spring-Summer 2010
Participating Customers - ECIPP – Site Reviews/On site Survey Gross and NTG	Census-attempt strata	30 interviews for 15 customers	NA	Impact Evaluation Team	Spring-Summer 2010
Participating Customers - ECIPP - Site Reviews/On site Survey Gross and NTG	~120 /year	45	90/10	Impact Evaluation Team	Spring-Summer 2010
Participating Customers - ECIPP - NTG Surveys	~120 /year *(same group as prior row)*	10	90/10	Survey Contractor	Spring-Summer 2010
Participating Vendors/ ESCOs - ECIPP - NTG Telephone Survey	>200	70	90/10	Survey Contractor	Summer 2010

Field work will span at least six months. The evaluators will plan to hold a meeting in early fall of 2010 after the major field work in the summer is complete to give prompt project design feedback to NYSERDA program staff and share preliminary findings with staff and DPS.

Key impact budget assumptions, especially those associated with the unit cost-related efforts and the telephone survey/interviews, are enumerated in Table 8. These constitute the major costs for data collection and constitute approximately two-thirds of the total costs for the proposed impact evaluation. The costs shown exclude costs for instrument development, data collection preparation and training,

further sample/population analyses, management, and reporting. These categories constitute the other third of the total cost for this program's impact evaluation.

Table 8. Impact Budget Basis

Budget Element	Cost	
Analysis - IM demand and energy savings interval data analysis (multiple IMs per site) 10% require verification-grade on-site inspection	$2,900	average per site
On-site visits and gross impact analysis for PDRE sample by M&V engineer Enhanced grade (50% of sites) Basic grade (33% of sites) Verification grade (17% of sites)	$8,900	average per site
Telephone interview with end-user for persistence study by technical expert 10% require verification-grade on-site inspection	$600	average per complete
Telephone interviews by technical expert for NTG	$450	average per complete

NYSERDA's program design ensures that larger EFP projects are subject to independent third party TA review as part of the implementation process. Verification-grade M&V evaluation alone therefore is expected to reveal little beyond persistence and might result in realization rates deceptively close to 1.0. This evaluation design thus invests in intensive high-quality site M&V to add maximum value in reducing uncertainty.

VII. Process Evaluation Plan

The next process evaluation of the EFP will be implemented in 2010. Process evaluations have been conducted for the former Peak Load Management Program, the Commercial & Industrial Performance Program and Enhanced Commercial & Industrial Performance Program. These process evaluations identified opportunities to improve the efficiency of the program processes as well as ways to work more effectively with ESCOs. The upcoming process evaluation will focus on the effectiveness of combining these programmatic efforts into a single point of entry for existing facilities to obtain support from NYSERDA. The process evaluation will include interviews with program staff, technical consultants supporting the program, ESCOs/vendors and end users who participate in the program.

Research Objectives

Anticipated research objectives for the 2010 process evaluation are listed below. In order for the process evaluation to provide the greatest value, other relevant or necessary objectives may be added, or objectives listed below may change somewhat, as the timing of the research draws closer.

1. Assess end user and ESCO/vendor experience with the program, specifically:
   a. Understanding of program requirements and program steps;
   b. Satisfaction with program services;
   c. Experience with the process for reviewing and approving projects and M&V requirements where applicable; and
   d. The decision making process for participating in the EFP, and how the EFP has affected the ESCO/vendor business.

2. Identify reasons for end user and ESCO/vendor participation in the program to assess:
   a. Awareness, knowledge and reasons for participation;
   b. Barriers to participation; and
   c. The decision making process for participation, as well as perceptions of the impact of external factors on participation (*e.g.*, financial climate, climate change, energy prices, etc.).
3. Document program progress and make recommendations for program improvement by assessing the:
   a. Program processes in providing appropriate services and quality assurance for different types and sizes of projects; and
   b. Process flow for the program services to assure they are efficient.

Activities

The technical consultant[25] and NYSERDA program staff interviews will be conducted early in 2010 to gather information about their perceptions of the EFP. The Process Evaluation Team will review the program database to identify projects at various stages of participation. While an average of approximately 300 projects are completed within the EFP each year, many more are in the program pipeline at any point in time. The process evaluation will sample projects from different stages in order to better understand the entire process of the program. Considering the variety of program services provided in the EFP, the process evaluation will also sample across the various program services. In early spring 2010, interviews will be conducted with a sample of ESCOs/vendors about current projects in the pipeline; concurrently interviews will also be conducted with end users associated with the sampled projects. The data collection will be completed by end of summer 2010 and a report will be prepared by the end of calendar year 2010.

Populations/Samples

Table 9 displays the samples assumed for direct work the Process Evaluation Team will implement. All key program staff and available technical consultants will be interviewed. The program database will be used to select a sample of participating projects, which will then be stratified by location (upstate/downstate), by stage of program participation, and by whether they are working with an ESCO or vendor. Past evaluations have shown that many of the ESCOs and vendors work on multiple projects; therefore the sample will be selected to include the range of ESCOs and vendors and project types represented in the current EFP. Given the variety of contractual arrangements observed in past evaluations, the total number of ESCOs and vendors active at the time the study will be conducted is difficult to predict.

Data Collection

Data collection for the process evaluation will begin with interviews with program staff and technical contractors in spring 2010. Interviews with staff will last about one hour, while those with technical consultants will last 30-45 minutes. Interviews with ESCOs and vendors will begin in late spring 2010 and continue into the summer. Conducted by the Process Evaluation Team, the ESCO/vendor interviews will last from 30-45 minutes. A survey of end users will be conducted during the summer of 2010 by NYSERDA's Survey Contractor Team. The end user surveys will last between 15 and 20 minutes.

[25] Technical consultants work under contract to NYSERDA to review project proposals and oversee the projects through quality assurance steps.

Table 9. Existing Facilities Program Process Evaluation Survey Specifics

Target Group	Estimated Population Size	Estimated Sample Size	Expected Sampling Precision	Administration By	Expected Start of Fielding
NYSERDA Staff	6	4	NA	Process Team	April 2010
Technical Consultants	10	8	NA	Process Team	April 2010
Energy Services Companies/Vendors	~200	50-68	90/10	Process Team	June 2010
End users	~600	100-136	90/10	Survey Team	June 2010

The process evaluation will focus on projects in the pipeline in 2010 and will therefore not be in direct conflict with efforts of the Impact Assessment Team to examine prior completed projects. Any information on decision making that might be of value to the Impact Assessment Team in their assessment of attribution will be passed along for consideration when the Impact Assessment Team conducts their evaluation of these projects.

Special Issues

Impact Evaluation - Ensuring No Real or Perceived Conflict of Interest

Megdal & Associates, NYSERDA's current Impact Evaluation Team, has four engineering firms, including ERS, GDS, Tow Path and Cx Associates, to perform the technical M&V services, providing the Impact Evaluation Team with the ability to meet schedule and efficiently cover a wide geographic range while still maintaining flexibility to avoid the potential for real or perceived conflict of interest. ERS currently provides quality assurance services to NYSERDA's Energy Efficiency Services group. Their detailed knowledge of NYSERDA programs, customers and program data is beneficial to the Impact Evaluation Team. The Impact Evaluation Team will avoid any real or perceived conflict of interest by assigning engineers to ensure that evaluation work is not undertaken by any individual that worked on the specific NYSERDA project in another capacity. The other three engineering firms have provided no implementation or technical assistance on NYSERDA-funded projects. None of the Impact Evaluation Team's engineering firms has directly provided equipment or performed installations for NYSERDA.

VIII. NYSERDA Evaluation Process

This evaluation plan is an early, but important step in NYSERDA's evaluation planning and implementation process. It is NYSERDA's understanding that DPS Staff wish to be involved as a reviewer/participant in the following parts of the evaluation process: detailed evaluation plans, project kick-off meetings, workplans (including sampling, statistics and modeling issues), data collection instruments, interim results reports (as applicable), presentation of evaluation results, and draft evaluation reports. NYSERDA will conduct evaluation planning and implementation in an open and transparent manner, and will invite DPS Staff participation in the designated aspects of the process and any others upon DPS' request.[26] Should DPS Staff choose to modify the level or manner of their involvement,

[26] In order to maintain transparency, and allow for confirmation checking and follow-up analysis, evaluation data will be maintained by NYSERDA and made available to DPS on an as-needed basis. NYSERDA will continue to maintain its secure "data warehouse" which includes data files, code books, and analysis files which can be made available in electronic form to DPS upon request. In order to provide a comprehensive record of each study conducted, the data warehouse also holds copies of final evaluation reports and appendices, including blank survey

NYSERDA should be notified about the change(s). DPS Staff should also choose when and how to involve their evaluation advisor consultant team in NYSERDA's evaluation processes, should directly provide any materials and information necessary for their advisor consultant team to fulfill this role, and should notify NYSERDA about the type and level of advisor consultant involvement.

An important goal of NYSERDA's evaluation effort is to provide early feedback to program staff to help inform and improve program implementation. NYSERDA accomplishes this goal in several ways:

1. Ongoing communications between the NYSERDA evaluation staff and evaluation contractors to identify issues that need to be brought to the attention of NYSERDA program staff, DPS Staff, and other involved parties.

2. Interim results reports may be generated, sometimes at the request of NYSERDA program staff and sometimes by initiative of NYSERDA's evaluation team and contractors, where early results are required or deemed useful prior to completion of the full evaluation effort.

3. Presentations of draft evaluation results held with NYSERDA evaluation contractors, evaluation team, program staff, and DPS Staff before evaluation reports are written provide feedback on the programs as soon as possible, and provide evaluation contractors with additional perspective and context that will be useful in reporting final recommendations.

Upon completion of final evaluation reports, the NYSERDA evaluation team will also provide support and assistance to program staff with regard to implementation of recommendations and program improvements.

IX. Reporting

Final reports will align with requirements set forth in the DPS evaluation guidelines, and will include: methodology, key results, recommendations, summary and conclusions, and appendices with detailed documentation.

Upon completion of each major evaluation study effort, findings and results will be communicated by NYSERDA's evaluation contractors and evaluation staff to NYSERDA program staff. Actionable recommendations and information on program progress toward goals will be provided as input to the program design and improvement process. NYSERDA's evaluation staff will follow up regularly with program staff on recommendations arising from the evaluation and the status of their consideration or adoption of these recommendations.

NYSERDA's evaluation staff will prepare quarterly and annual reports to the Public Service Commission, DPS and the EAG summarizing the results of all programs and from all evaluation studies occurring in the most recent quarter or year. The latest evaluated program savings, realization rates, and net-to-gross ratios will be used in compiling data for these overarching reports. Quarterly reports will be provided to the Commission within 60 days of the end of each calendar quarter. The annual report will substitute for the fourth quarterly report, summarizing program and portfolio progress throughout the calendar year. The annual report will be submitted to the Commission within 90 days of the end of the calendar year.

---

instruments, although these documents will be made available to DPS and publicly upon completion of each evaluation project.

X. Total Resource Cost Analysis

Once per year, NYSERDA will update benefit/cost ratios (at a minimum, Total Resource Cost test) for each major program and for the entire portfolio of SBC-funded **New York Energy $mart**[SM] and EEPS programs. The Total Resource Cost (TRC) test divides the present value of the benefits by the present value of program and participant costs. A benefit-cost ratio greater than 1 indicates benefits exceed NYSERDA and participant costs. The Program Administrator Cost (PAC) test divides the present value of the benefits by the present value of the program administrator costs. A benefit-cost ratio greater than 1 indicates benefits exceed NYSERDA costs. For more detailed definitions of benefit/cost terms and a description of NYSERDA's current benefit/cost input sources, including avoided energy, capacity and distribution costs, refer to Appendix A of NYSERDA's September 22, 2008 Energy Efficiency Portfolio Standard Program Administrator Proposal. The latest evaluated program savings, realization rates, and net-to-gross ratios resulting from the evaluation efforts described in this plan will be used in the annual benefit/cost analysis update. If available, NYSERDA will also present benefit/cost scenarios that include non-energy impacts.

NYSERDA will conduct benefit/cost analysis for its programs in a manner consistent with other program administrators, as appropriate. NYSERDA has knowledgeable staff and tools in place to accomplish benefit/cost analyses for all of its SBC and EEPS programs. NYSERDA is prepared to make adjustments to its current practice should DPS Staff or the EAG decide that alternative methods, tools, or inputs are superior or would foster greater consistency among program administrators.