

# **NYSERDA FLEXIBLE TECHNICAL ASSISTANCE PROGRAM**

## **Evaluation, Measurement, and Verification Plan**

July 16, 2009

### **I. Introduction**

The detailed evaluation plan presented in this document builds upon the brief evaluation plan included in NYSERDA's August 22, 2008 filing of the *System Benefits Charge Supplemental Revision for New York Energy \$mart<sup>SM</sup> Programs (2008-2011)* for the Flexible Technical Assistance (Flex Tech) Program. In order to revise and add detail to its original evaluation plan submittal, 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 7<sup>th</sup>, 2008 and to provide the highest level of rigor possible within the available resources.

As the Flex Tech Program ramps up to meet the aggressive EEPS 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 Flex Tech Program which is supported by SBC and EEPS funding. NYSERDA will not differentiate between funding sources when conducting this evaluation effort.

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

The overarching goals of NYSERDA's *New York Energy \$mart<sup>SM</sup>* 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), Department of Public Service (DPS) staff, and other stakeholders with timely and unbiased information regarding program implementation. Specifically, the goals for the Flex Tech Program evaluation are to:

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

- (3) Track changes in markets over time with a specific focus on market indicators that are likely to be impacted by the Flex Tech Program;
- (4) Assess customer satisfaction of the quality and timeliness of program services and contractors
- (5) Examine the internal program processes, interaction with other implementation programs, as well as public awareness of the program and steps to participate.
- (6) Assess reasons for and barriers to participation, and explore decision-making processes related to participation

The Flex Tech Program budget (3<sup>rd</sup> Quarter 2008 through 2011) consists of approximately \$14.9 million in EEPS funds and \$29.4 million in SBC funding, providing a total budget of \$44.3 million. The proposed evaluation budget is \$2.3 million which equates to approximately 5% of program funding.<sup>1</sup> Evaluation budgets are detailed in Table 1.<sup>2</sup>

**Table 1. Flex Tech Evaluation Schedule and Budget**

Evaluation Element	Estimated Budget and Completion						% of Total Evaluation Budget
	2009	2010	2011	2012	2013	Total	
Market Characterization & Assessment	\$182,000a,b	-	\$183,000a,b	-	-	\$365,000	15%
Impact Evaluation	\$757,000	\$53,000	\$488,400	\$31,000	\$543,860	\$1,873,260c	80%
Process Evaluation	-	\$107,000	-		-	\$107,000d	5%
<b>Total</b>	<b>\$939,000</b>	<b>\$160,000</b>	<b>\$671,400</b>	<b>\$31,000</b>	<b>\$543,860</b>	<b>\$2,345,260</b>	<b>100%</b>

a. Primary data collection costs represent approximately 40% of the total proposed evaluation budgets.  
b. The two evaluations are scheduled for the latter months of the years presented above (*i.e.*, the budget shown in 2009 will carry-over into 2010 and the budget shown in 2011 will carry-over into 2012).  
c. This includes SBC III and EEPS evaluation funding. The 2009 impact evaluation is for the 2007-2008 program, but counts toward New York’s 15X15 goal. Approximately two thirds of the total costs are for data collection.  
d. The process evaluation costs include \$12,000 for interviews, and \$25,000 for surveys.

### III. Flex Tech Program Description and Goals

The Flexible Technical (FlexTech) Program provides customers with objective and customized information to facilitate informed energy efficiency, procurement, productivity and financing decisions. Cost-shared technical assistance is provided for detailed studies from energy engineers and other experts. The program is designed to evaluate all energy sources while providing objective analysis of energy source trade-offs and switching options. Program participants receive a customized energy study targeted to the participant’s particular needs and objectives.

<sup>1</sup> 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.

<sup>2</sup> The evaluation budget is roughly three times the amount proposed in the 60-day Fast Track plan due to the addition of SBC program funds which are approximately double the EEPS funds allocated to the Flex Tech Program.

Eligible participants for the FlexTech Program include commercial, industrial, institutional, municipal, not-for-profits and K-12 schools. Participants may use NYSERDA-contracted or customer-selected service providers. The FlexTech Program is currently offered Statewide with an increased focus on the Consolidated Edison service territory due to load constraints, higher energy costs and ratepayer contributions. This geographically-focused application of the FlexTech Program targets service providers located in the Consolidated Edison area and offers higher cost sharing limits for customers in this service territory.

Smaller customers are currently eligible for walk-through energy audits, including a reimbursement of audit cost upon implementation of recommendations. NYSERDA anticipates that targeting of this market sector will be diminished as the utility-offered Fast Track Small Business Programs begin to be implemented.

Projected MWh savings for the FlexTech Program are shown in Table 2.

**Table 2. Projected MWh Savings for FlexTech Program (2008-2015)**

	2008 (4 <sup>th</sup> quarter)	2009	2010	2011	2012	2013	2014	2015	TOTAL
EEPS	0	5,888	23,515	49,360	67,230	62,789	38,480	19,848	267,109
SBC III	0	12,406	34,117	60,258	70,449	59,372	33,674	17,723	287,999
<b>TOTAL</b>	<b>0</b>	<b>18,294</b>	<b>57,632</b>	<b>109,618</b>	<b>136,679</b>	<b>122,161</b>	<b>72,154</b>	<b>37,571</b>	<b>555,108</b>

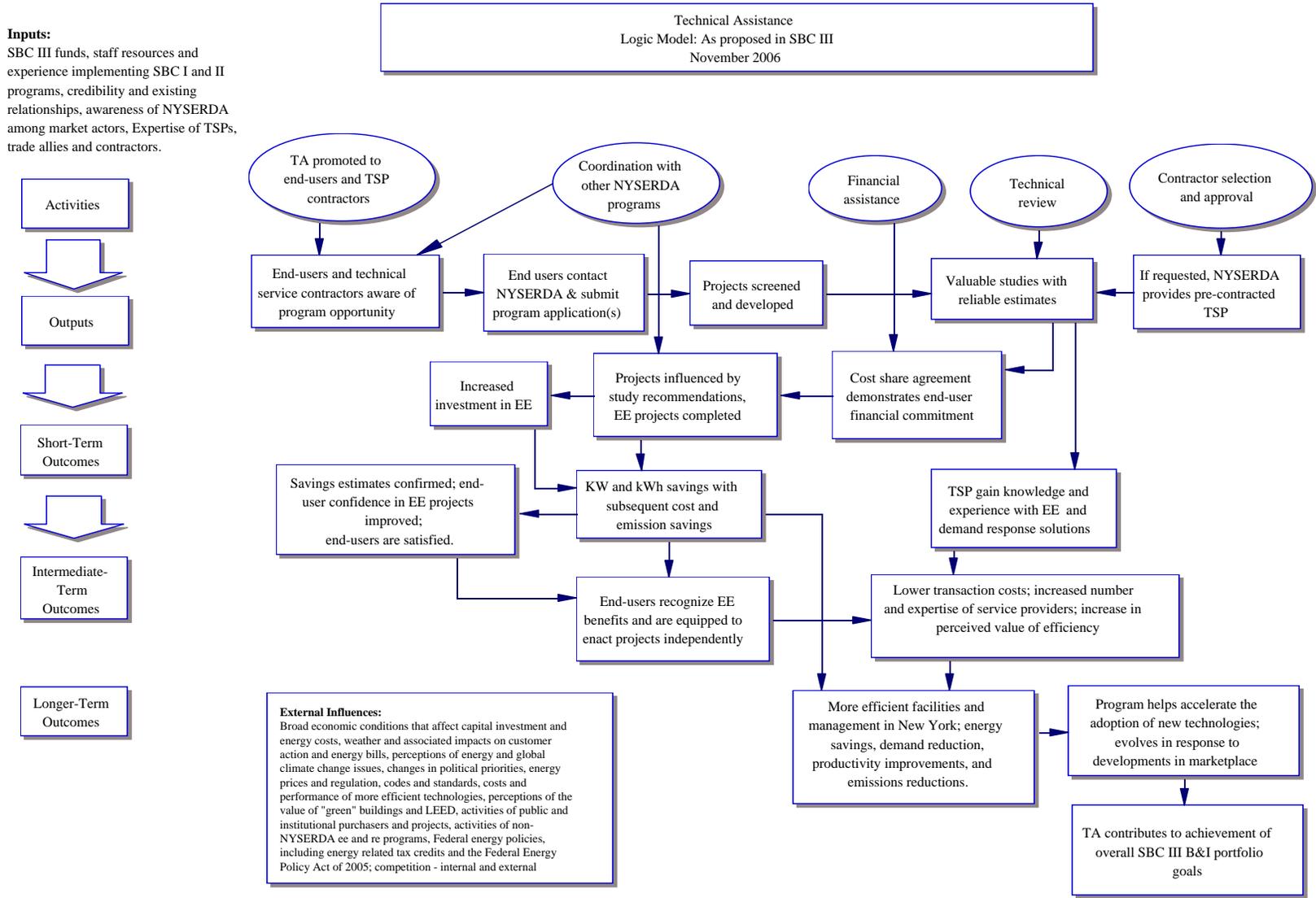
Source: NYSERDA, *System Benefits Charge Supplemental Revision for New York Energy Smart<sup>SM</sup> Programs (2008-2011)* As Amended, August 22, 2008 and Revised March 12, 2009, Page 7, Table 2-3.

#### **IV. Logic Model/Theory**

Figure 1 presents the most recent logic model for the Flex Tech Program. As program evaluation efforts begin, a first step in the process will be to review the latest logic model and make updates to the model 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. Flex Tech Logic Model**



## V. Market Characterization & Assessment Plan

This section presents the Market Characterization and Assessment (MCA) evaluation plan for the Flexible Technical Assistance (Flex Tech) Program.

### *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 Flex Tech Program 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 under which NYSERDA operates 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*

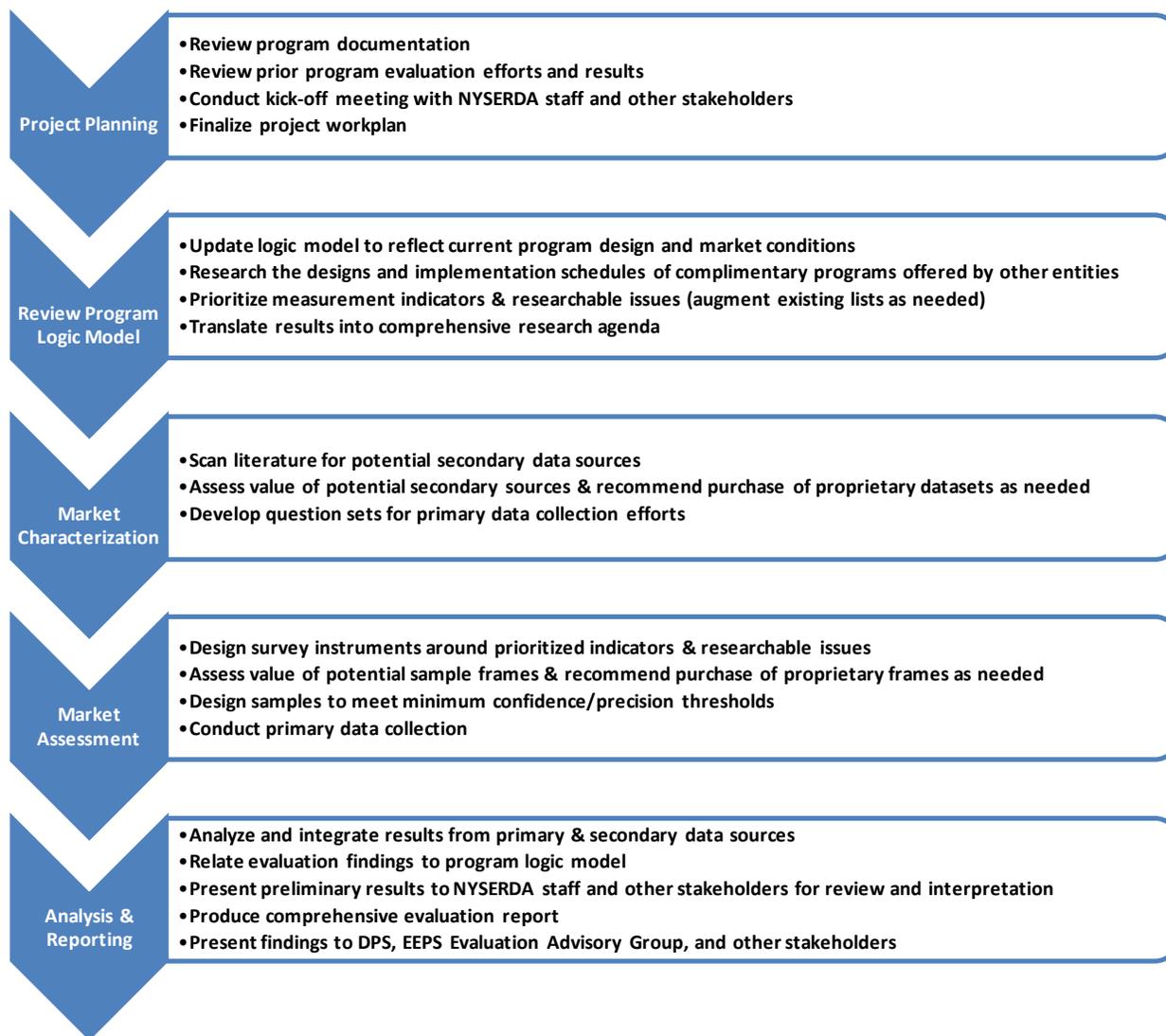
The proposed MCA evaluation plan for the Flex Tech Program consists of multiple activities (blue arrows) and associated research tasks (bulleted lists), as shown in Figure 2. The approach will make use of a variety of primary and secondary data sources to generate information on a number of topics relevant to the Flex Tech Program including: program accomplishments and market share in terms of both equipment replacement activity and interaction with key market actor groups; changes in customer and technical service provider/contractor awareness and understanding of efficiency measures and practices promoted by the program; and customer motivations and decision-making criteria related to the upgrade or retrofit of energy-using systems in existing nonresidential buildings. The approach is driven primarily by elements and theories presented in the Flex Tech *Program Logic Model Report*<sup>3</sup>, and key research findings generated by the evaluation will be related to the outputs and outcomes anticipated by the program logic model. In addition, the approach is intended to encourage a high degree of interaction between the MCA Team and NYSERDA program and evaluation staff as well as DPS staff and other project stakeholders via project planning activities and deliverable review cycles. The MCA Team welcomes active engagement by these parties but is cognizant of the possibility that other demands may limit the parties' contributions during certain points in the evaluation process.

Each activity and the associated research tasks are discussed in more detail in the remainder of this section.

---

<sup>3</sup> NYSERDA, *Technical Assistance Program – Program Logic Model Report*, November, 2006. See Section III of this document for additional details regarding the FlexTech Program Logic Model.

**Figure 2. Synopsis of Flex Tech 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 Flex Tech Program 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 Flex Tech Program 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.).

## Review Program Logic Model

The Flex Tech *Program Logic Model Report* was designed to help guide NYSERDA's program-specific evaluation activities; thus, an initial activity undertaken by the MCA Team will be to conduct a comprehensive review of the *Program Logic Model Report* to ensure the document accurately reflects the current program design and state of the market. A central element of the review will be researching the designs and implementation schedules of complementary energy efficiency programs being administered by utilities and other parties (both within New York and in other states) to identify potential leveraging opportunities wherein NYSERDA and the other program administrators can possibly collaborate to achieve broader and deeper program impacts. The results of this review, including the MCA Team's suggested prioritization of measurement indicators and researchable issues, will be presented to NYSERDA staff in memorandum format and suggested updates to the document, if any, will be discussed with NYSERDA staff and other project stakeholders to reach consensus on the proposed revisions.

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 logic model review 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 Flex Tech 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) and Manufacturing Energy Consumption Survey (MECS) data, U.S. Census County Business Patterns Reports, membership lists and other publicly-available data from relevant professional organizations (*e.g.*, the Building Owners and Managers Association (BOMA), the International Facility Management Association (IFMA), etc.), 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:

- Existing building stock segmented by market sector, geography, and energy use patterns
- Firmographic information regarding customers and technical service providers/contractors participating and not participating (*i.e.*, remaining market potential) in the Flex Tech Program
- Prevailing supply chains, business cycles, and technical service delivery channels
- Flex Tech Program accomplishments and market penetration in terms of both statewide equipment retrofit/replacement activity and interaction with existing technical service provider networks
- Other metrics as identified

## Market Assessment

Market Assessment results will be generated through primary data collection efforts with end-use customers and technical service providers/contractors participating in the Flex Tech Program 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 during the logic model review.<sup>4</sup> 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 perceptions regarding value of technical assistance services promoted by the program (*e.g.*, energy feasibility studies, rate analysis and aggregation studies, retro-commissioning, and long-term energy management)
- Market awareness of NYSERDA program offerings and broader energy efficiency opportunities
- Customer decision-making processes in terms of engaging technical assistance service providers and making equipment replacement decisions
- Capability of market infrastructure to provide technical assistance services and structure of relationships between technical assistance service providers and their customers
- Technical service provider/contractor expertise with energy efficiency measures and services including emerging technologies/designs as well as LEED<sup>®</sup>/green building and renewable energy options
- 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 Flex Tech Program 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, Flex Tech Program 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

---

<sup>4</sup> Other evaluation contractors will be able to suggest additions to the instruments to collect data relevant to separate studies and the MCA Team will endeavor to accommodate such requests balancing the additional survey components against the need to minimize impacts on survey respondents.

deemed final by NYSERDA staff and other project stakeholders. Final evaluation results will also be presented to DPS staff and other project stakeholders during scheduled meetings.

### ***Populations/Samples***

As discussed previously, the MCA evaluation of the Flex Tech Program will involve primary data collection with end-use customers and technical service providers/contractors participating in the program as well as with comparison non-participant groups eligible to participate in the program.<sup>5</sup> The Flex Tech Program tracking database will be the sample frame for program participants. Participants eligible to be included in the sample frame will be those end-use customers and technical service providers/contractors associated with projects that had audit reports completed subsequent to April 30, 2006.<sup>6</sup> Based on the findings from the Market Characterization work, the participant sample frame will be stratified to reflect meaningful differences in program participation. For participating end-use customers, stratification variables might include project size (*i.e.*, expected kWh savings), energy efficiency technology, or market sector. For participating technical service providers/contractors, stratification variables might include contractor type (*e.g.*, ESCO vs. trade ally) and area of expertise (*e.g.*, lighting, HVAC systems, motors, compressed air). The final stratification plan will be designed to meet the 90/10 absolute confidence/precision criteria for all participants on an upstate-downstate regional basis.<sup>7</sup>

All non-residential buildings and most commercial and industrial organizations are eligible to participate in the Flex Tech Program. However, it would be prohibitively expensive to develop a comprehensive sample frame representative of that population. Moreover, since analysis of program participation should identify those market sectors for which the Flex Tech Program offers the greatest energy saving potential, it would be appropriate to focus the non-participant data collection on two groups:

- **Active Market Sectors** – The analysis of the program database will furnish information on the market sectors that are most active in the Flex Tech Program. By interviewing non-participants in these same market sectors, it will be possible to assess additional market potential for the Flex Tech Program within the sectors.
- **Inactive Market Sectors** – The discussions with Flex Tech Program staff and the analysis of the program database will allow the MCA Team to identify market sectors that appear to have substantial energy savings opportunities, but that have limited participation to date in the Flex Tech Program. By interviewing non-participants in inactive market sectors, it will be possible to measure the level of awareness and understanding of the Flex Tech program in inactive sectors and thereby assess the additional program potential that might be achieved by targeting new markets.

---

<sup>5</sup> 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.

<sup>6</sup> Projects that had audit reports completed prior to April 30, 2006 were sampled during previous MCA studies.

<sup>7</sup> 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.

Once the target market sectors (both Active and Inactive) have been identified, the MCA Team will work closely with NYSERDA's data collection contractor, APPRISE, to identify potential sample frames for non-participating end-use customers and technical service providers/contractors and develop sampling procedures to effectively represent the targeted sectors. The most efficient approach to the sample frame development will be to identify potential list frames for each market sector and to use a hierarchal list frame development procedure that will maximize the coverage of the eligible populations, while also ensuring the efficiency of sample frame development. The sample allocation will be designed to furnish 90/10 absolute confidence/precision on an upstate-downstate regional basis.<sup>8</sup>

Current estimates regarding estimated sample sizes, expected sampling precision, and anticipated survey fielding dates for the 2009 MCA evaluation are summarized in Table 3.<sup>9</sup> These estimates will be finalized prior to undertaking the planned evaluation and once the MCA Team more thoroughly analyzes program participation data.

---

<sup>8</sup> 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.

<sup>9</sup> 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.

**Table 3. Flex Tech Program MCA 2009 Evaluation Specifics**

Target Group	Estimated Population Size	Estimated Sample Size	Expected Sampling Precision <sup>1</sup>	Survey Administration By	Expected Start of Fielding
Participating End-use Customers	TBD	140a	90/7	Survey Contractor	Summer 2009
Each Region (Upstate/Downstate)	TBD	70	90/10	Survey Contractor	Summer 2009
Participating Technical Service Providers/Contractors	TBD	140a	90/7	Survey Contractor	Summer 2009
Each Region (Upstate/Downstate)	TBD	70	90/10	Survey Contractor	Summer 2009
Non-participating End-use Customers	TBD	140a	90/7	Survey Contractor	Summer 2009
Each Region (Upstate/Downstate)	TBD	70	90/10	Survey Contractor	Summer 2009
Non-participating Technical Service Providers/Contractors	TBD	140a	90/7	Survey Contractor	Summer 2009
Each Region (Upstate/Downstate)	TBD	70	90/10	Survey Contractor	Summer 2009

<sup>1</sup> Assumes proportional sampling, two-tailed test, finite population correction

<sup>a</sup> 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.

The MCA Team will be conducting research during the 2009/2010 timeframe on a number of different programs targeting Commercial and Industrial organizations. For each program, the research will include an analysis of program participation. After comparing participant groups among the different programs, the MCA Team may find that there is substantial overlap among the organizations participating in the different programs. If so, the team will design an end-use customer survey covering those programs that are substantially overlapping to minimize respondent burden and reduce design and implementation costs. Similarly, the team may also find that there is substantial overlap in the contractors participating in the different programs. If so, the team will design a contractor survey covering those programs that are substantially overlapping.

**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<sup>10</sup> and will consist of the following steps undertaken by the survey contractor: 1) format the final survey instruments and program them into a CATI system, 2) pretest the final instruments with subsets of the market actor

<sup>10</sup> Surveys will be designed to be completed in approximately 15 – 20 minutes.

group samples and consult with the MCA Team as needed to resolve any issues that are identified<sup>11</sup>, 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 MC&A 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. Doing so 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 VI). The Flex Tech Program 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 Flex Tech Program are shown in Table 4. These initial budget estimates will be finalized prior to undertaking the planned evaluation after participant sample sizes are determined through analysis of program participation data.

**Table 4. Flex Tech Program MCA Evaluation Schedule and Budget**

Evaluation Element	Estimated Budget and Completion					
	2009	2010	2011	2012	2013	Total
Market Characterization & Assessment	\$182,000a,b	-	\$183,000a,b,c	-	-	\$365,000
a Primary data collection costs represent approximately 40% of the total proposed evaluation budgets. b The two evaluations are scheduled for the latter months of the years presented above ( <i>i.e.</i> , the budget shown in 2009 will carry-over into 2010 and the budget shown in 2011 will carry-over into 2012). c Funding to support the 2011 MCA evaluation may come out of future evaluation budgets if the Flex Tech Program is continued beyond 2011.						

## VI. Impact Evaluation Plan

Historically, the Flex Tech Program has generated between 75 and 150 project reports each year, although that number will increase significantly with the higher funding levels. SBC III generated savings from the Flex Tech to-date are over 174 GWh,<sup>12</sup> with one-third of those downstate and two-thirds upstate. The 2009 evaluation will evaluate Flex Tech SBC-funded projects completed in 2007 and 2008. The first impact evaluation of the EEPS-funded effort will be conducted in 2011, allowing time for project completions from 2009 and the first half of 2010. The final EEPS evaluation will be conducted in 2013 evaluating all Flex Tech projects completed in 2010 and 2011.

<sup>11</sup> Pretest interviews will be included as completed interviews unless major revisions to the instruments are made.

<sup>12</sup> These are cumulative program savings since July 1, 2006.

## ***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 energy, demand, and MMBtu savings, which will entail verifying the installation and the estimation of savings for a representative sample of program participants for comparison to an established baseline or possibly to a matched sample of non-participants. The gross savings realization rate is then applied to the program population or NYSERDA-reported gross savings to derive the final savings estimates (evaluation-based estimates, or *ex post* savings).

Since the primary function of the Flex Tech Program is to provide detailed studies with recommendations for efficiency improvements to participating firms rather than offering direct incentives, an additional component of the gross savings impacts is determining the number of participants who followed through and installed the recommended measures (the measure adoption rate or MAR).

### 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, the net-to-gross ratio (NTGR) becomes the adjustment factor to derive net impacts.

For the Flex Tech Program, net effects will be measured for participating companies and the firms that provide the technical assistance reports. This assessment of net effects will cover participant spillover. Non-participant spillover from Flex Tech could easily overlap with NYSERDA's other program targeted toward the Commercial/Industrial (C/I) sector, and these effects are planned to be measured through a study of the entire C/I existing facilities market scheduled for 2009.

### Precision and Bias

Sample sizes will be designed to target 90/10 precision at the program level. Large sample sizes are planned for estimating the MAR, since telephone surveys are relatively inexpensive, and it is possible that the results for the MAR for all years may meet the 90/10 confidence/precision standard within each utility territory. However, given the high expense associated with estimating the savings realization rate (SSR), the Impact Evaluation Team has concluded that it will not be feasible to add the additional sample size required to meet the 90/10 standard at the regional (upstate/downstate) or

utility level for the overall gross savings estimates. 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 Flex Tech Measurement & Verification (M&V) methodology is designed to address the unique nature of the program, whereby NYSERDA shares the cost of analyzing and reporting on energy and demand savings opportunities for a facility. No action is required by the facility and no incentives (through this program) are offered for improvements.

The realization rate has two components.

- The first component is the measure adoption rate (MAR). It reflects the percentage of savings from measures recommended in completed studies that have been implemented in whole or in part. Evaluators will determine the MAR based on a telephone survey of a large sample of participants. Surveys will be conducted on this sample repeatedly over the 2009-2010 period following each project through completion or until the outcome for each recommendation at the sampled site is definitively resolved (*i.e.*, implemented or will definitely not implement). As part of this survey evaluators will also track which of the implemented measures received NYSERDA, utility, federal, or other funding assistance, and which measures are not implemented but planned for near-term adoption.
- The second realization rate component is the savings realization rate (SRR). This is the percentage of estimated savings achieved by the implemented measures. For identified and implemented measures funded under Tier 3 of the Enhanced Commercial/Industrial Performance Program (ECIPP), the savings realization rate associated with that program will be applied to the measure.<sup>13</sup> For a sample of the measures implemented but not funded by ECIPP, an evaluation engineer will conduct verification site visits for those sites with substantial savings. A sub-sample of site visits will be subject to monitoring and/or measurement, meeting as a minimum the standards of IPMVP Option A including the use of direct measurement. Sites in the largest savings stratum may be subject to IPMVP Option B level analysis.<sup>14</sup> Savings will be estimated, using engineering models, based on reported baseline conditions (or code assumptions) and as-built conditions.

Impact evaluations were conducted on the Flex Tech/Technical Assistance (TA) program for NYSERDA by Nexant in 2004 and 2006.<sup>15</sup> These prior studies strongly suggest that program participants use the TA reports to make changes over time (perhaps to coordinate with other updates that are occurring). The 2006 study found MARs of 29% shortly after the study was received, a MAR of 55% after one year and over 70% four or more years after participation. Surveys of participants from four or more years since participation obtained a high response rate of almost 80%. The 2009 evaluation will build upon what has been learned in these earlier evaluations and the data collected in order to create more reliable estimates of the MAR over time.

---

<sup>13</sup> The realization rate for ECIPP will be applied to measures recommended in the technical assistance study and then installed through the ECIPP program. This method is justified due to the rigorous level of M&V that the ECIPP program applies to each project.

<sup>14</sup> The criteria for “large savers” will be determined after review of the program database and results of the initial MAR surveys are available for review.

<sup>15</sup> The 2004 study included surveys to develop the MAR, file review and M&V site visits. The 2006 study included surveys and updates to the MAR and file review.

The initial MAR sample, as further described below, will be drawn from participants historically through 2008. The first MAR survey will attempt to contact this full sample. Additional follow-up calls will be placed to this sample every six to twelve months to follow-up on whether additional recommendations have been adopted.<sup>16</sup> This compares to earlier studies where MAR update calls had been conducted every 2 years. Increasing the frequency of calls improves recall and is more likely to reach those familiar with the original technical assistance study. Customers will be dropped from the sample as all of their recommendations for that site have been definitively resolved. Additional incremental sample for the TA participants from 2009-2010 will be added to the MAR survey sample in 2011. Another incremental sample addition will occur in 2013 for Flex Tech participants from 2011. MAR calls will be occurring every six to twelve months from mid-2009 through 2013. The MAR evaluation is a cumulative evaluation across evaluation years (a rolling sample), with the most recent MAR findings being reported in each program evaluation report.

Several of the commercial and industrial programs within the NYSERDA portfolio are designed to work together to serve the energy efficiency needs of New York ratepayers by addressing different barriers that hinder customer adoption of efficient practices and technologies. Some programs supply information, others promote product availability, and many directly pay down the incremental cost of more efficient but more expensive technologies.

The very synergies that NYSERDA has promoted and which are successfully motivating rate payers to increase their efficient use of energy result in a significant reporting challenge: how to attribute the savings from a project to a program. If a building owner has used technical assistance funds to conduct a feasibility study, and then implemented some or all of the recommendations using financing available through the Loan Fund, and has also received rebates from Smart Equipment Choices (one of the predecessors of the Existing Facilities Program), then the savings will be tracked in all of these programs. This allows programs to get full credit for any savings they had influenced and for the program benefit/cost estimates to not be under-estimated.

NYSERDA has recognized this issue of overlapping savings and addressed it by discounting the combined savings at the sector and portfolio level before publishing energy and demand savings data. The evaluation study was completed in September 2008, measuring the proportion of overlapping savings from overlapping projects with the Existing Facilities program (and its predecessor programs CIPP, ECIPP, and SEC<sup>17</sup>) and the Distributed Generation and Combined Heat and Power program (DG/CHP). The savings from TA overlap with Existing Facilities was found to be 19.3%. Overlapping savings with DG/CHP was 26.8% for energy savings and 20.0% for demand savings. These overlap estimates are used by NYSERDA in reporting correct sector level and portfolio level savings.

The SRR is a combination of the applicable ECIPP realization rate for those measures installed through ECIPP and additional evaluation engineering performed through this evaluation. Measurement and verification (M&V) will involve site visits and engineering modeling of installed measures for the participating customers in the sample who have installed measures, as well as an analysis of energy use data (both electric and gas). On-site metering or measurements will be incorporated into the M&V plans as needed. Utility bills will be obtained and used to calibrate the savings estimate, where possible.

---

<sup>16</sup> The Impact Evaluation Team will examine the prior data, have discussions with staff, and experiment with the optimal follow-up time period. This plan assumes a nine-month follow-up period.

<sup>17</sup> CIPP, ECIPP, and SEC are the Commercial/ Industrial Performance Program, the Enhanced CIPP and the Smart Equipment Choice programs, respectively.

Evaluators recognize that it will be challenging to engage Flex Tech participants in site visits and measurement of implemented measures that NYSERDA did not fund. Several options to help meet this challenge are already under discussion.

### 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), *i.e.*, without the study.<sup>18</sup> 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, due to the study, additional actions are taken to reduce energy use at the same project site, but these actions had not been included in the technical assistance study.
- Outside project spillover occurs when a participant in the program initiates additional actions that reduce energy use at other sites and were not part of the technical assistance study or any other NYSERDA program, but were caused by what the participant learned through the technical assistance study or the implementation of its recommendations at the participating site.

In addition, non-participants can also be influenced by the program. The influence of NYSERDA’s Flex Tech program on the C/I sectors can easily overlap with the influence of the NYSERDA’s other major commercial and industrial non-new construction program, the Existing Facilities program. Recognizing this, NYSERDA conducted a C/I non-participant spillover study applicable across C/I programs in 2005 and 2007. NYSERDA plans to conduct a similar but expanded study in 2009 to derive updated non-participant spillover rates for both the Flex Tech program and the Existing Facilities program. (The 2007 C/I non-participant spillover rate from this prior study will be used until the 2009 update has been completed.) Consequently, this component of the Flex Tech NTG effects will not be addressed in this section.

The Impact Evaluation Team intends to explore participant free-ridership and spillover for the Flex Tech Program with participating customers and vendors through an enhanced self-report survey process. This is another example of the enhancements afforded by additional funding. The decision-making process will be investigated from both viewpoints to determine the most reliable way to combine free-ridership survey responses from these two groups. Other decision-makers, such as chief financial officers or vendors, will also be interviewed for the largest projects, if they are found to be heavily influential in the decisions to invest in energy efficiency measures.

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 is complicated as we are asking about conjecture of a theoretical alternative. Use of 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

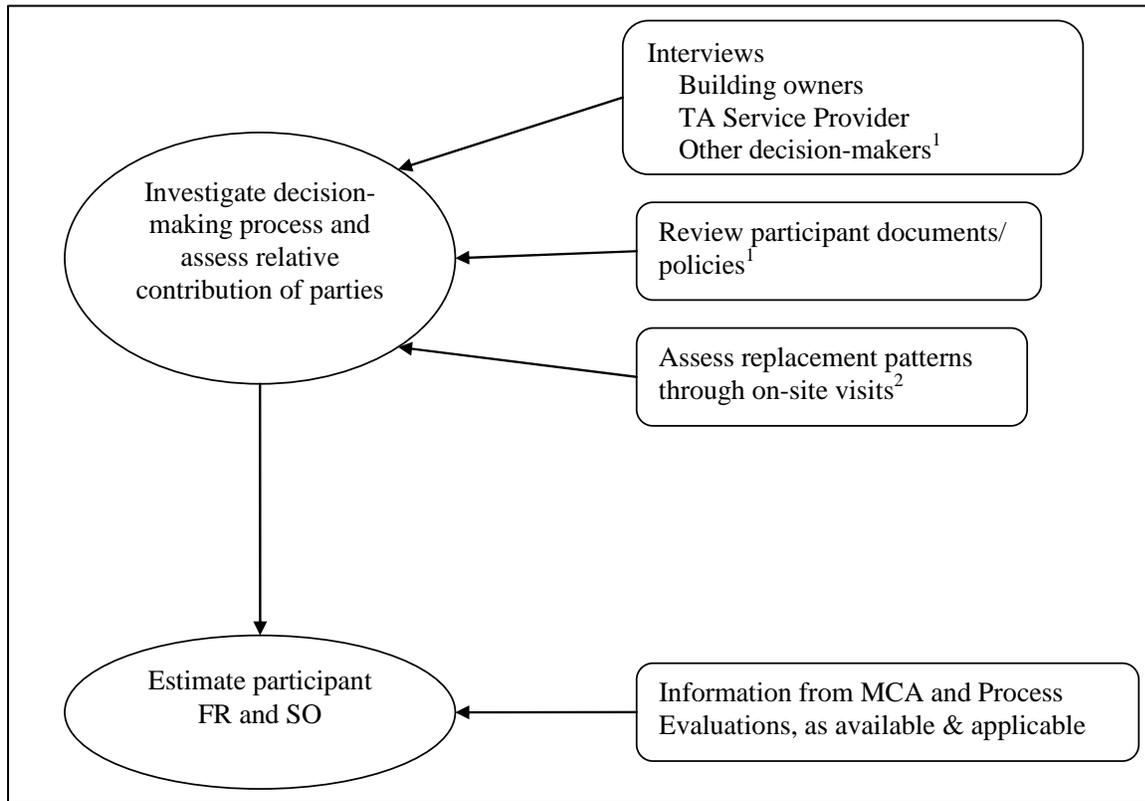
---

<sup>18</sup> Note that savings accrued due to the technical assistance study then funded by either the customer or a NYSERDA implementation program are accounted for under gross savings.

estimate. Measuring free-ridership in multiple ways can increase the construct validity of the estimate.

If the statewide baseline studies being proposed by NYSERDA are undertaken, these results will be analyzed and compared to the Flex Tech participant sample to derive a market-based estimate as 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 results will also be compared against the participant enhanced self-reports described above. These alternative methods will be combined to estimate a draft, triangulated net-to-gross (NTG) ratio (NTGR) that will provide a high level of construct validity for the NTG estimates. 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 TA free-ridership and participant spillover estimates. The enhanced self-report components and overall process for the development of these estimates is illustrated in Figure 3.

**Figure 3. Enhanced Self-Report Process for the TA Program’s Participant FR & SO Estimates**



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 implementation will be conducted to minimize self-selection to the extent possible. Obtaining cooperation and good response rates from customers for the NTG surveys can be especially difficult with customers that may have only received incentives for one-half of the study costs and then made the entire efficiency investment on their own. The Impact Evaluation Team and NYSERDA staff will explore alternative methods to address this challenge in 2009, with some preliminary improvements to be tested in the 2009 evaluation and a more refined plan for the later Flex Tech evaluations.

### ***Populations/Samples***

Sampling will be necessary to estimate both gross and net impacts, as discussed in more detail below.

### **Gross Impact Sampling**

The first component of the verification of gross savings is the telephone surveys designed to determine whether measures were actually installed at the site. In addition, the planned impact evaluation will include significant site survey work on a representative sample of participants. The sampling will be nested in that the results from the telephone survey, *i.e.*, the list of completed projects with measures installed, will be used as the sample frame for the on-site survey.

The sample frame for the telephone survey will be all firms who received technical assistance reports after a specified cutoff date (possibly January 1, 2002) and have not yet installed all recommended efficiency measures. The purpose of establishing a cutoff date is to avoid sampling aging projects that are unlikely to have followed the recommendations and also to limit the evaluation to those projects that can reasonably be expected to have been affected by the NYSERDA technical assistance study. Prior NYSERDA TA evaluation research found that customers can take several years to install recommended measures. The developed TA adoption rate curve does not plateau for as long as seven years after study completion, suggesting that the cutoff date should not be more recent than 2002.

Data collected for the previous Nexant evaluation will be reviewed as appropriate to identify firms that are known to have installed some or all of the recommendations and the sample frame will be adjusted accordingly. Firms that are known to have installed all of the recommended measures will be removed from the sample frame and firms that installed only some of the measures may be included in the sample frame to determine whether the remaining recommendations are still under consideration.

The Impact Evaluation Team will consider whether the phone survey should be stratified to ensure representation in the sample for firms receiving the technical assistance study in each year subsequent to the cutoff date. Stratifying by the timing of the report would ensure that the MAR rates would be based on a representative sample of older and more recent projects. The Impact Evaluation Team estimates that 300 firms will receive at least one telephone survey and some may be surveyed on multiple occasions until at least some measures are installed or it is concluded no implementation will occur.<sup>19</sup> MAR surveys with this sample will occur every six to twelve months through 2009 and 2010. Assuming a nine-month (averaging the six- to twelve-month options) follow-up and resolution rate of 50 customers per follow-up period, there will be a total of 650 follow-up calls in addition to the original 300 calls through 2010.<sup>20</sup> This sample size is expected to be sufficient to estimate the MAR for electric and gas measures at the program-level within the 90/10 confidence/precision standard. This is a large advance over past M&V studies. Due to the increase in available funds, NYSERDA is able to increase confidence and precision levels from 80/20 by using larger sample sizes.

A rolling sample will be used with the MAR surveys to determine overall MARs (measure adoption rates). An incremental sample of 100 will be added to the MAR survey sample in 2011 from the Flex Tech participants of 2009 and 2010. Another incremental sample of 100 will be added in 2013 from the 2011 Flex Tech participants.

The greatest uncertainty for this program's gross savings estimate is associated with the measure adoption rate (MAR), *i.e.*, the rate at which the recommended measures were installed. Large sample sizes are planned for estimating the MAR (as described above). This should allow the MAR estimate for all years to meet the 90/10 confidence/precision level downstate and upstate. It may also be possible to obtain this sampling precision level for the MAR within each utility territory.

---

<sup>19</sup> The sample size of 300 is based on the assumption that there are 600 projects with technical assistance reports provided from January 1, 2004 through the end of 2008, and that separate MAR rates will be developed for gas and electric measures. Using a coefficient of variance of 1.0, assuming a target confidence/precision level of 90/10 for both categories of projects and incorporating the finite population correction factor gives a sample size of 200 projects for gas and 200 for electric. The Impact Team assumes that 100 projects can overlap and be used for both the electric and gas samples, and 100 projects will need to be selected separately for each of the two categories, resulting in a projected sample size of 300.

<sup>20</sup> The 650 follow up calls is based on the assumption that an average of 2.2 calls per project will be made within the time frame of the 2009 evaluation cycle.

For the on-site survey, efficient sample sizes will be chosen using stratified ratio estimation (SRE) to meet a 90/10 confidence/precision level for the statewide program over the entire evaluation cycle. The sample size necessary to meet this precision target statewide is estimated to be approximately 65 projects for program years 2007 and 2008. This sample size is based on a SRE sample with an estimated error rate of 0.6 and incorporating the finite population correction factor. Given the high costs of conducting on-site verification, the sample for the SSR will not be designed to meet the 90/10 confidence/precision target upstate/downstate or at the utility level.

Projects will be stratified by size (typically the magnitude of the energy savings), timing (year of the report), type of measure (*e.g.*, on-site generation v efficiency measures), or other variables, as indicated. A census of large energy-saving sites and a sample (meeting 90/10 confidence/ precision levels) of remaining sites in each stratum will be selected for verification site visits. The smallest savers may be eliminated as site visit candidates.

### Attribution Sampling

The evaluation of net impacts is focused on participating customers and vendors. The TA providers selected for the NTG survey will be those matched with the sample of participants, starting with the site visit sample and supplemented to reach the required sample sizes needed for a goal of 90/10 sampling precision both upstate and downstate for the NTG factor estimates.<sup>21</sup> The Impact Evaluation Team will consider whether this project-centric sampling approach will provide a sufficient sample of vendors. When both the participating customers and the engineering firms/vendors 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 Evaluation 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. The Impact Evaluation Team estimates that approximately 70 unique engineering firms/vendors will be associated with the 130 projects selected for the NTG surveys.

### *Data Collection*

To be able to conduct the sampling and proceed with the evaluation, the Impact Evaluation Team will need the following information from the NYSERDA Flex Tech staff at a minimum:

- Project level information, including address, contact information for the site owner and engineer, the type of project (custom, design/build), type of business
- Measure level information (in easily readable electronic format), such as a description of the measure, quantity recommended, the energy savings (electric, gas and other fuels), demand savings, measure life, installation costs
- Firmographics, including the size of the firms, the number of employees, the fuels used for major end uses, types of major electric and gas end uses

In addition, critical information will need to be collected from third party sources, as described below.

---

<sup>21</sup> In general, the decision-makers for the participating firms and those best able to assist with the technical implementation questions undertaken within the MAR survey are two different individuals.

- Utility consumption data (both electricity and natural gas) for participants, 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, weather station
- Weather data, which may be available from the utilities or from the national weather service

Evaluators will collect original data as described in the Activities above.

Table 5 displays the target audiences, sample sizes and schedule for surveying in 2009.

**Table 5. FlexTech 2009 Impact Evaluation Survey Specifics**

Target Group	Estimated Population Size	Estimated Sample Size	Expected Sampling Precision	Survey Administration By	Expected Start of Fielding
Participating Firms \ MAR Telephone Survey	~75 to 150/year	300	90/10	Impact Evaluation Team	Fall 2009
Participating Firms - On-Site Survey	~75 to 150/year	65	90/10	Impact Evaluation Team	Late Fall 2009
Participating customers - NTG decision-maker surveys	~75 to 150/year	130 projects <sup>1</sup>	90/10 Upstate & Downstate	Survey Contractor	Fall 2009
Matched participating TA service providers (& supplemental sample to reach desired sample size)	35-40/year	70	90/10 Upstate & Downstate (on project basis)	Survey Contractor	Fall 2009
<sup>1</sup> Customers surveys are assumed to total 150 for the 130 due to interviewing multiple decision-makers for the largest projects. The incremental cost for any additional customer surveys is \$100 per complete.					

Key impact budget assumptions, especially those associated with the unit cost-related efforts, are enumerated in Table 6. Unit costs constitute approximately two-thirds of the total costs for the proposed impact evaluation. The costs shown exclude fixed costs for instrument development, data collection preparation and training, database preparation and data collection system development, further sample/population analyses, management, and reporting. These categories constitute the other one third of the total cost for program evaluation.

**Table 6: Flex Tech Impact Budget Basis**

Budget Element	Cost
Medium or large, performance-based project with pre/post metering - Verification or Option A level M&V (75% of sites) - Basic or enhanced post-retrofit M&V (25% of sites)	\$2,100 average per site \$10,300 average per site
Telephone interviews for MAR: file review and initial MAR interview & data (assumes an average of 5 hours per customer: initial contact, initial interview, MAR data, analysis to MAR database entry)	\$575 per customer
Telephone interviews for MAR: follow-up interviews with previously called customers (assumes an average of 1 hour per customer: calls to contact, MAR data, analysis to MAR database entry)	\$115 per customer
NTG Telephone Survey	\$100 per complete

Unit costs include escalation for the later-year budgets to best approximate the costs to be incurred at that time. Fixed costs such as those associated with design, reporting, and management include both escalation and discounts recognizing that replicating a study is less expensive than executing it for the first time.

**Schedule and Budget**

Table 7 outlines the Impact budget by year for the Flex Tech Program.

**Table 7. Flex Tech Impact Evaluation Budget**

Evaluation Element	Estimated Budget and Completion					
	2009	2010	2011	2012	2013	Total
Impact Evaluation	\$757,000	\$53,000	\$488,400	\$31,000	\$543,860	\$1,873,260a
a. This includes SBC III and EEPS evaluation funding. The 2009 impact evaluation is for the 2007-2008 programs, prior to EEPS but counting towards New York 15X15 goals.						

**VII. Process Evaluation Plan**

The process evaluation of the FlexTech Technical Assistance for Existing Buildings Program builds on the one previous process evaluation of the Technical Assistance (TA) program and the TA logic model. The process evaluation will be primarily formative and will examine the experience of participants in the pipeline for the program, from application stage up to four or six months after receipt of a study.

The evaluation will rely on interviews with the program implementation team and surveys of participants. Data collection will occur in spring of 2010 focusing on projects in the application pipeline in 2009. Depending on whether program funding continues beyond 2011, a second process evaluation could be conducted to address any ongoing issues, to follow-up on findings of the 2010

evaluation and to explore any new issues that are identified at that time. This second data collection period could focus on TA reports in the pipeline in 2011 and could occur in 2012, providing approximately 18-24 months between data collection periods.

The process evaluation will also work closely with the Market Characterization and Assessment (MCA) team to include process evaluation questions in the surveys of nonparticipating existing building owners and technical assistance providers.

### ***Research Objectives***

The following are the primary process evaluation objectives. 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 this research draws closer.

1. Examination of program processes including but not limited to:
  - a. Assess program process flow and opportunities for streamlining
  - b. Assess the coordination processes for the program with other programs
  - c. Understand sources of program awareness, knowledge of efficiency opportunities and perceptions of program steps of participation
  
2. Identify reasons for participation in the program including but not limited to:
  - a. Assess reasons for participation and reasons for partial participation
  - b. Assess barriers to participation, perceptions of importance of energy efficiency
  - c. Assess decision making process for participation, perceived impact of external factors on participation (financial climate, energy prices, climate change)
  
3. Identify decision making processes regarding measure implementation including but not limited to:
  - a. Assess expectations for implementation of measures, key issues to consider in decision making process, anticipated time frames for implementation
  - b. Assess barriers to implementation, perceived impact of external factors on decision, perception of value of energy efficiency investment
  - c. Assess awareness of other programs that could assist in implementation, assess perception of these programs
  
4. Document program progress and make recommendations for program improvements
  - a. Assess the effectiveness and efficiency of the program in increasing technical assistance capability and capacity for existing buildings
  - b. Assess satisfaction with the quality and timeliness of FlexTech services
  - c. Assess customer satisfaction with program experience, perceptions of cost-share agreement, perception of quality of program services

Key issues for the second process evaluation will be identified at that time. Issues that may be important include follow up on issues identified in the first study, exploration of program attrition, and continuing assessment of decision-making processes as the landscape for energy efficiency evolves.

### ***Activities***

The process evaluation will begin with meetings with program staff and a review of program database and program documents to identify participants and their stage of participation. After the population of participants has been identified, the process team will conduct interviews with program staff and Flex Tech contractors to document their experience of the program and the program processes and to develop a clear understanding of the program process flow.

The process evaluation team will select a sample of participants from the program database. The population will include a full range of participants from those that only have an application submitted to those that have had a study completed for four to six months, thus including active as well as partial-participants who may have dropped out of the program process. Surveys will be conducted with the sample of participants and partial participants and with customer-selected technical assistance providers for those participants included in the survey.

***Populations/Samples***

Five populations that are the focus of data collection as shown in Table 8. The process evaluation team will interview the program manager and staff involved in working with existing building customers as well as a sample of existing building Flex Tech contractors. A sample of participating and partial participating existing building customers will be selected based on the analysis of the database, the sample will be stratified based on location (upstate/downstate), project kWh size as well as status in the program pipeline. The team will also interview any customer selected technical assistance providers associated with sampled participant and partial participant projects.

***Data Collection***

The review of the program database for projects in the pipeline in 2009 will begin in early 2010. As shown in Table 8, the process evaluation team will conduct interviews with program staff and FlexTech contractors in early spring 2010. The survey of participants and partial-participants will begin in summer 2010 and will likely be completed by early fall 2010. During this same period interviews will be conducted with customer-selected technical assistance providers associated with selected customer projects. Analysis and report writing will be completed by early by December 2010. Interviews with program staff and Flex Tech contractors will last about one hour. Interviews with the customer selected technical assistance providers will last about 20 minutes. The surveys with participants will last 15-20 minutes and with partial participants 10 minutes.

**Table 8. FlexTech Process Evaluation Survey Specifics for 2009-2010**

Target Group	Estimated Population Size	Estimated Sample Size	Expected Sampling Precision	Survey Administration By	Expected Start of Fielding
NYSERDA Program Staff	10	4	NA	Process Team	March 2010
FlexTech Contractors	42	21	80/10a	Process Team	March 2010
Customer Selected Technical Assistance Providers	100	20-30	TBD	Process Team	June 2010
Participants	500	134	90/10b	Survey Contractor	June 2010
Partial-Participants	100	56	90/10b	Survey Contractor	June 2010
a Assumes proportional sampling, 2-tailed test, finite population correction, absolute precision.					
b Assumes proportional sampling, 2-tailed test, absolute precision.					

### *Special Issues*

To support the M&V evaluation, the customers who are surveyed will be notified of their possible inclusion in an additional study. The process team will provide process questions for inclusion in the MCA survey of nonparticipating existing building owners and technical assistance firms.

### *Schedule and Budget*

Table 9 displays the schedule and budget allocation for the process evaluation. The process evaluation budget estimate includes data collection costs of \$12,000 for interviews, and \$25,000 for surveys.

**Table 9. FlexTech Evaluation Schedule and Budget**

Evaluation Element	Estimated Budget and Completion					
	2009	2010	2011	2012	2013	Total
Process Evaluation	-	\$107,000	-	-	-	\$107,000

### **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.<sup>22</sup> Should DPS Staff choose to modify the level or manner of their involvement, 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.

---

<sup>22</sup> 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 instruments, although these documents will be made available to DPS and publicly upon completion of each evaluation project.

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, finding 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 on 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.

## **VIII. 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<sup>SM</sup> 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 definition 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.

NYSERDA will conduct benefit/cost analysis for its programs in a manner consistent with other program administrators, as appropriate. NYSERDA has knowledgeable staff and a tool 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.