IMPACT EVALUATION

NYSERDA 2007 – 2009 FlexTech Program

FINAL Report

Prepared for

The New York State
Energy Research and Development Authority

Rebecca Reed and Emily Shusas
Project Managers

Prepared by
Megdal & Associates Impact Evaluation Team
Satyen Moray, Principal Investigator, ERS
Jonathan B. Maxwell, PE, ERS
With Team Members
Lori Megdal, Megdal & Associates
Kathryn Parlin, West Hill Energy & Computing

Project Number 9836

NYSERDA
March 2012

NOTICE

This report was prepared by a team led by Megdal & Associates in the course of performing work contracted for and sponsored by the New York State Energy Research and Development Authority (hereinafter the Sponsor"). The opinions expressed in this report do not necessarily reflect those of the Sponsor or the State of New York, and reference to any specific product, service, process, or method does not constitute an implied or expressed recommendation or endorsement of it. Further, the Sponsor, the State of New York, and the contractor make no warranties or representations, expressed or implied, as to the fitness for particular purpose or merchantability of any product, apparatus, or service, or the usefulness, completeness, or accuracy of any processes, methods, or other information contained, described, disclosed, or referred to in this report. The Sponsor, the State of New York, and the contractor make no representation that the use of any product, apparatus, process, method, or other information will not infringe on privately owned rights and will assume no liability for any loss, injury, or damage resulting from, or occurring in connection with, the use of information constrained, described, disclosed, or referred to in this report.

ABSTRACT

NYSERDA's Flexible Technical Assistance Program funded through the System Benefits Charge (SBC) Operating Plans (FlexTech) provides cost-sharing for objective and customized energy efficiency information (energy studies) to commercial, industrial, institutional, government, and not-for-profit customers. This report describes an impact evaluation of the FlexTech Program completed studies for the calendar years 2007 through 2009.

Engineers conducted a telephone survey of facility managers to determine the measure adoption rate (MAR) and date of adoption for measures recommended in 303 studies completed between January 1, 2003 and September 30, 2009. Engineers visited forty-four participants and completed site-specific measurement to assess the proportion of savings being realized compared to the study-recommended savings recorded by NYSERDA. Survey professionals interviewed participants and service providers that conducted the FlexTech work to assess free ridership and participant inside and outside spillover effects.

The presented results of the Program include estimates of the measure adoption as of the time of the survey; the expected year-by-year and final term measure adoption rate overall, by fuel type and by measure technology; savings realization rates of installed measures by fuel type; and free ridership, spillover, and net-to-gross factor. The report concludes with recommendations regarding future program operation and future evaluation activities.

ACKNOWLEDGMENTS

This report was prepared with input from Rebecca Reed, Judeen Byrne, and Emily Shusas of NYSERDA. The Impact Evaluation Team would like to also thank Brian Platt and Jaime Ritchey.

CONTENTS

APPENDICES

EXECUTIVE SUMMARY

NYSERDA's Flexible Technical Assistance Program (FlexTech) is funded through the System Benefits Charge (SBC) Operating Plans. It provides objective and customized energy efficiency information to commercial, industrial, institutional, government, and not-for-profit customers. The goal of the FlexTech Program is to increase productivity and economic competitiveness of participating facilities by identifying and encouraging the implementation of cost-effective energy efficiency, peak-load curtailment, combined heat & power (CHP), and renewable generation studies.

The Program is designed to evaluate energy-related opportunities and the majority of a study must address electrical impacts. Program participants receive cost-shared analysis targeting their particular energy and business needs.

Participants may use either their own independent consultants or FlexTech consultants under NYSERDA contract that are competitively selected to provide a statewide geographic distribution of needed technical services.

EVALUATION OBJECTIVES

The primary purpose of this FlexTech impact evaluation was to estimate the savings attributable to the Program. The evaluation needed to account for multiple layers of decision-making and performance in order to estimate the installed savings resulting from study funding, specifically: (1) what recommended measures study recipients implemented, (2) how much those implemented measures save compared to the study's predictions, (3) how NYSERDA's study funding influenced the recipients' decisions both to conduct the study and to implement measures, and (4) how participant experience influenced them to take additional actions beyond those recommended in the studies.

To achieve this goal of estimating attributable savings, the evaluation team determined:

1. **Measure adoption rate (MAR) -** A ratio that quantifies the percentage of study-recommended savings that customers chose to adopt. This factor solely addresses decision-making. It does not consider evaluated performance.
2. **Savings realization rate (SRR)** –A ratio that divides the evaluated savings for installed measures by the study-recommended savings for installed measures. The product of the MAR and SRR is the overall realization rate.
3. **Free ridership (FR)** –The likelihood the participant would have received the study or implemented the measure without the Program.
4. **Inside participant spillover (SO) -** The degree to which the end users' participation in the FlexTech Program influenced them to take additional actions to save energy.
5. **Outside participant spillover (SO)** – The degree to which vendors' participation in the Program influenced their practices with other non-program customers.
6. **Nonparticipant spillover (SO)** – The degree to which the Program influenced non-participating end users and vendors to save energy through measures not funded by the Program.

The research evaluated the impact of studies completed from January 1, 2007 through September 30, 2009. The MAR survey sample was selected to estimate the MAR for electric and natural gas measures at the study level and for upstate and downstate within the 90/10 confidence/relative precision standard. For the on-site survey, sample sizes were chosen using stratified ratio estimation (SRE) to target a 90/10 confidence/relative precision level for the statewide program over the entire evaluation cycle. The net-to-

gross (NTG) survey was a census attempt of decision-makers for adopting sites. Non-participant spillover (SO) and other indirect program effects are within the purview of a separate study.

Some of the implemented measures received implementation funding from NYSERDA or other sources. Determining the amount of this programmatic overlap falls within the scope of a separate study and is not part of this impact evaluation. NYSERDA applies this overlap factor at the portfolio level.

As secondary objectives, however, this evaluation collected survey data regarding program funding overlap to aid later research as well as collecting information on the age of replaced equipment to help NYSERDA learn more about early replacement decisions.

RESEARCH APPROACH

The evaluation team estimated three factors (MAR, SRR, and NTG) to arrive at the final savings attributable to the FlexTech Program, referred to as the program net impact. The three factors were developed from data collected in the following series of surveys:

1. **MAR survey -** Engineers conducted a telephone survey of facility managers or engineers to determine the MAR and date of adoption for measures recommended in 303[1] studies completed between January 1, 2003 and September 30, 2009.
2. **SRR survey -** Engineers led an on-site survey with measurement and verification (M&V) to estimate the SRR for adopted measures associated with forty-four studies completed between January 1, 2006 and September 30, 2009 and for which the MAR survey identified at least one adopted measure. Two large combined heat and power (CHP) studies were also included in the on-site survey[2]. The site-specific M&V was a mixture of "basic" and "enhanced" levels of rigor data collection and analysis.
3. **Participating end user NTG survey -** Survey professionals conducted a telephone interview-based survey of decision-makers from forty-seven sites associated with study recipient organizations to aid in determining FR and SO. This sample was a subset of the MAR survey respondents that received studies after January 1, 2006 and had adopted at least one measure.
4. **Service provider NTG survey -** Survey professionals administered a telephone interview-based survey with FlexTech service providers concerning forty-six sites. The participating FlexTech service providers all were associated with studies in the decision-maker sample.
5. **Additional decision-maker NTG survey -** Researchers attempted to interview additional decision-makers in the NTG sample's census stratum. Ultimately, there was only one study for which two additional decision-makers were identified and interviews completed.
6. **On-site observation and policy survey -** The engineers that conducted the site M&V also made observations regarding equipment purchasing practices and spoke with the on-site contact about efficiency-related purchasing policies while on-site for SRR work.

[1] Researchers completed telephone interviews associated with 301 Flextech studies. There were two additional studies for which the MAR results were already known, for a total of 303 completed MAR questionnaires.

[2] These two studies were not included in the MAR survey because the installations were known to have been completed.

The NTG survey instruments and data collection were designed to replicate the method used in the prior program evaluation and to also calculate NTG factors on an enhanced[3] basis using data collected in the last two surveys above. While the additional decision-maker survey was ultimately ineffective, the on-site observation and policy was informative and was used to adjust upper and lower FR bounds as determined from the participating end user and service provider NTG surveys.

Table 1-1 below presents the general timeline of when the survey activities were conducted.

Table 1-1. Survey Periods

Survey Type	Date Range
MAR telephone survey	May 2010 to July 2010
On-site survey	December 2010 to March 2011
NTG telephone surveys	March 2011 to May 2011

FINDINGS

The following sections present the overall results developed by the evaluation team associated with the MAR[4], SRR, and NTG estimates.

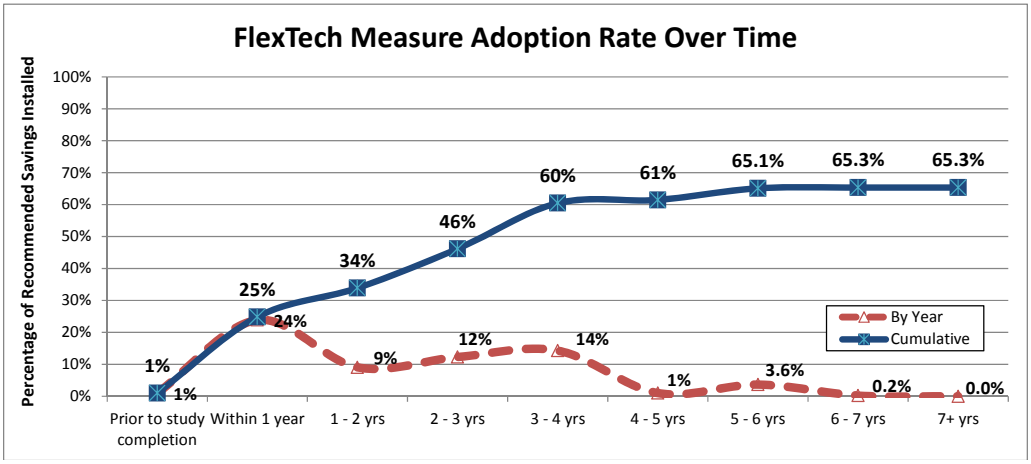
Measure Adoption Rate Results

Figure 1-1illustrates the MAR over time, based on customer responses associated with 2,452 measures recommended in the 303 studies. Eventually 65% of the savings associated with recommended measures was implemented by study recipients. Just over two-thirds of that savings was realized within 3 years of the study completion, but some implementation continues through the 6th year.

[3] The "enhanced" NTG methods included two added elements: (1) interviews of multiple end-user decision-makers per study instead of just one and (2) direct on-site observation by engineers regarding equipment replacement and capital investment practices.

[4] Two rounds of MAR telephone surveys were conducted during the span of this evaluation. The MAR results presented in his report are from the more extensive round 1MAR telephone survey. The round 2 MAR telephone surveys were fielded a year after (May 2011) the round 1 MAR telephone surveys to update the implementation status on measures that were unresolved during the round 1 MAR telephone survey. Due to the timing of the round 2 MAR telephone survey, its results could not be incorporated into this study. The round 2 MAR telephone survey resulted in a very minor revision to the round 1 MAR. Additional details about the round 2 MAR telephone survey can be found in Appendix M at the end of this report.

Figure 1-1. Program MAR Over Time



The MAR plateaus at different adoption rate percentages depending on measure and fuel type. Table 1-2 summarizes the projected long-term eventual MARs for electric and non-electric efficiency measures and on-site generation. Non-electric measures have the lowest adoption rate, electric efficiency is just above the 0.65 mean, and generation measures are adopted at a slightly higher rate than efficiency.

Table 1-2. FlexTech Projected Long-Term Measure Adoption Rates

Measure Type	Long-Term Projected Measure Adoption Rate
Electric energy efficiency	0.67
Non-electric energy efficiency	0.42
On-site generation	0.72
Overall	0.65

Table 1-3 presents the MAR results as of the summer of 2010 by the study completion year and by geographic region. The MAR for electric measures meets the 90/10 confidence/precision target for the Program as a whole as specified in the evaluation plan. The variability in the natural gas savings is due to the combination of efficiency measures (with savings) and generation projects (with extra use), and also the relative low value of the MAR. The relative precision for natural gas energy efficiency measures alone is 8.5%.[5]

[5] The sample was designed to achieve 90/10 on electricity and natural gas. The upstate natural gas relative precision was driven unexpectedly upward due to the low MAR of 0.275. If it had been about 0.50, as the other MARs were, relative precision would have met the 10% relative precision goal.

Table 1-3. Measure Adoption Rates at Time of Survey and Sampling Statistics

Parameter	Electric Energy			Natural Gas			On-Site Generation Electricity Only	Total Excluding Natural Gas Generation
	Upstate	Downstate	Total	Upstate	Downstate	Total		
MAR as of Summer 2010	0.64	0.60	0.63	0.26	0.60	0.31	0.72	0.56
Number of studies in frame	506	149	655	298	60	358	41	655
Total sample	236	67	303	141	25	166	31	3037
Standard error	0.017	0.026	0.014	0.029	0.004	0.025	0.015	0.014
Relative precision at 90% confidence	4.07%	6.64%	3.48%	17.58%	1.03%	12.44%	3.15%	3.73%
Coefficient of variation	0.380	0.330	0.369	1.269	0.031	0.975	0.106	0.485

Using the collected information, we calculated the MAR as a function of elapsed time since study completion. The findings of the MAR survey are as follows:

- The mean time to adopt was 1.5 years, and measures that saved 70% of the eventual total savings were installed within 3 years of study completion, but measures continue to be adopted even in the 6th year after study completion before they plateau.
- Most CHP studies require 2 to 3 years to implement. Two large CHP studies required 3 and 6 years for completion.
- Within a specific technology group, controls savings were most frequently adopted by study recipients, with a savings adoption rate 25% greater than that for the next most readily adopted technology, which was lighting.
- Controls measures were also the most often recommended, representing 25% of the time for non-generation measures. This controls trend is an interesting finding, as lighting often is perceived as the most common opportunity and most readily adopted due to low uncertainty and lack of complexity. The pattern can likely be explained by the relatively low cost and fast payback time of many controls measures, but it also indicates customer willingness to implement measures that tend to be more complex.
- Envelope measures were the least frequently adopted.
- On average, customers adopted more of the larger savings measures and fewer of the lesser savings measures.
- Comparison of the results between this study and the prior one shows both similar patterns of adoption over time and similar results when methodological differences are minimized to the extent possible. This study found slightly higher adoption rates overall.

Savings Realization Rate Results

Table 1-4 and Table 1-5 present the overall SRR results summary for electricity and natural gas. The Program had an overall realization rate for annual electric energy savings of 0.92, with 3.7% relative precision.

The Program's reported demand savings are based on a previous M&V program evaluation completed by Nexant[6]; they are not based on site-specific information in the studies. Evaluated savings are based on site-specific analysis. The results demonstrate that the factor produced reasonable estimates for the sample; the associated SRR is 0.99. The factor is not as applicable for continuous operation CHP studies. The demand realization rate overall is 0.86. While relative precision estimates accompany the demand realization rates, they are marginally informative given the method the Program used to estimate demand.

[6] Nexant, *M&V Evaluation Technical Assistance Program*, prepared for NYSERDA, June 2007.

Table 1-4. SRR Results Summary, Electricity

Studies	Total Number of Studies	Number of Completed Sites in SRR Sample	Electric Energy		Summer Coincident Peak Demand	
			SRR[1]	SRR Relative Precision	SRR	SRR Relative Precision
Upstate	106	37	0.94	6.3%	0.80	29.9%
Downstate	15	7	0.87	2.9%	1.05	0.6%
Total	121	44	0.92	3.7%	0.86	12.5%

[1]SRR includes 1.0 SRR for forty-one small studies not in sample (constituting less than 3% total energy savings).

Table 1-5. SRR Results Summary, Natural Gas Efficiency

Region	Natural Gas Efficiency	
	SRR	SRR Relative Precision
Upstate	0.77	4.0%
Downstate	0.80	0.0%
Total	0.77	4.0%

SBC reports and tracking systems are for all-fuel MMBTU. This involves significant amounts of fuel besides natural gas for CHP fuel energy use. The CHP combined non-electric SRR is 0.89.

The evaluators found that operating differences, calculation methods, and analysis parameter differences were the largest contributors to the deviations observed between the evaluation and FlexTech study results.

Net-to-Gross (NTG) Results

The sample frame for the NTG surveys was the same as the sample frame for the on-site survey, i.e., studies with at least one installed recommended measure and FlexTech reports provided to the participant from January 1, 2006 to September 30, 2009. The NTG surveys were a census attempt of the end users and the TA Service Providers for the adopting sites. Table 1-6 below lists the FR, inside SO, and outside SO estimates developed for the current evaluation.

Table 1-6. Current Free Ridership, Inside Spillover, and Outside Spillover Estimates

Attribution Variable	Factor
Free ridership	0.32
Inside spillover	0.04
Outside spillover	0.30
Non-participant spillover	0.15
Net-to-gross factor	1.17

The current FR estimate is close to the prior estimate developed by MCAC[7] in 2007 on the 2005-2006 studies, which was 27%[8].

The current SO rates are high but within typically acceptable range. The prior inside SO estimate was 11% and outside SO estimate was 19%.

The non-participant SO uses the same prior non-participant SO study with NPSO at 15%.

Therefore, the overall NTG factor is: 1 - 0.32 + 0.04 + 0.30 + 0.15 = 1.17.

The prior NTG factor was 1.17.

Net Impact

Table 1-7 presents the eventual expected impact of FlexTech studies completed in 2007 through 2009.[9]

Table 1-7. Net Long-Term Expected Impact, Studies Completed 2007 - 2009

Parameter	Sample Frame Period	Electric Energy (MWh/yr)	Electric Demand (MW)	Natural Gas (MMBtu/yr)
Study recommended and tracked savings	1/1/07 – 12/31/09	146,651	24.6	928,023
MAR, long-term expected	1/1/03 – 9/30/09	0.68	0.68	0.43
SRR	1/1/06 – 9/30/09	0.92	0.86	0.77
Adjusted gross impact		91,745	14.4	307,268
NTG	1/1/06 – 9/30/09	1.17	1.17	1.17
Net impact, long-term expected		107,342	16.8	359,504

Table 1-8 presents the overall net impact as of the summer of 2010 for studies completed in 2007 through 2009. The main difference between Table 7 and Table 8 is the MAR. The MAR in Table 7 is based on responses of participants from 2003 through 2009 and includes projected savings from the 2007-2009 studies' measures expected to be implemented after evaluation field work was completed in the summer of 2010. The MAR in Table 8 is based solely on participants from 2006 through 2009 and only accounts for savings from measures already installed as of the time of the evaluation.

[7] *Technical Assistance Program Market Characterization, Market Assessment and Causality Evaluation, Final Report*, Prepared by Quantec LLC and Summit Blue Consulting, LLC, 2007.

[8] The FR rate and net-to-gross ratio (NTGR) that NYSERDA uses in its quarterly and annual reports is a blended rate over the previous evaluations and used to match the fact that NYSERDA reports cumulative savings. The comparison reported here is strictly between the prior causality evaluation and this impact evaluation to ensure a valid comparison.

[9] The MARs in this table differ slightly from those of Tables 1-2 and 1-3 because the prior table's values separate generation from efficiency, whereas the MAR for overall net impact by fuel is blind to measure technology type. SRRs are calculated excluding the negative impact measures.

Table 1-8. Net Impact as of Summer 2010, Studies Completed 2007 - 2009

Parameter	Electric Energy (MWh/yr)	Electric Demand (MW)	Natural Gas (MMBtu/yr)
Study recommended and tracked savings	146,651	24.6	928,023
MAR as of summer 2010	0.35	0.35	0.31
SRR	0.92	0.86	0.77
Adjusted gross impact	47,222	7.4	221,519
NTG	1.17	1.17	1.17
Net impact as of summer 2010	55,250	8.7	259,177

MAR, SRR, and NTG Summary

Table 1-9 summarizes the MAR, SRR, and NTG factors found in this evaluation.

Table 1-9. FlexTech Projected Long-Term MAR, SRR and NTG

Studies	MAR by Measure Type[1]		SRR by Energy Parameter			NTG
	Electric Energy	Fossil Fuels	Electric Energy	Summer Coincident Peak Demand	Fossil Fuels	All
Efficiency	0.68	0.43	0.86	0.82	0.77	1.17
CHP	0.72		1.00	0.88	0.89	

[1] This table shows a single value indicating the long-term MAR for each category. Actual data entry will be associated with the year-by-year curves that plateau at the shown MARs.

RECOMMENDATIONS

The principal goal of the current evaluation effort was to analyze the impact of the FlexTech Program studies completed in 2007-2009. The evaluation team also observed opportunities to improve operation and savings estimation in the future to hopefully narrow the variation in realization rates. Key recommendations include the following:

Program Recommendations

- **Focused marketing of controls studies/vendors**– The MAR survey results indicated that controls measures had the highest adoption rate of all technology groups. Aggressive promotion of this particular type of study could increase the overall cost-effectiveness of the Program by increasing the MAR. Additional investigation regarding why controls measures have such a high adoption rate could lead to lessons learned, which could be applied to studies associated with other technologies.
- **FlexTech study data post-processing** - Consider adding a step to the final NYSERDA database (buildings portal) data entry. A third-party contractor or the service provider could be hired to review the FlexTech study findings and translate the results into data needed for reporting requirements and impact evaluations. In many cases no translation will be necessary but for some it will help to:
  - Consistently report the peak demand savings. NYSERDA reports the average demand savings during summer weekday afternoons to the New York Department of Public Service (NY DPS). NYSERDA's tracking system should retain demand savings estimates on the

same basis. The FlexTech studies variously reported demand savings as the average demand savings over 12 months, the sum of demand savings for each of the 12 months, the annual super-peak demand savings, zero (omitted entirely when it existed), etc. Hence the evaluators found considerable variability in the reported realization rates associated with the peak electrical demand savings.

- Capture interactive savings associated with central cooling and heating plants. Some studies focused on energy savings in individual buildings within a large campus, where the buildings were served by a central boiler, chiller, or CHP plant. The studies reported savings in Btus of steam or hot water, ton-hours of chilled water, or campus distributed electricity to the building according to the interests of the recipient. In some cases the sub-metered savings was omitted entirely from NYSERDA's tracking system (ton-hours), understating overall study savings, and in other cases the tracking system noted the sub-metered energy impact instead of the impact at the customer's utility meter, potentially distorting NYSERDA's eventual benefit-cost and emissions calculations. NYSERDA should track the net savings at the customer's utility meter, not at the building Btu sub-meters, even if the sub-meter is the parameter of interest to the study recipient. The evaluators recommend that either the service providers be asked to provide such results or that NYSERDA calculate them. This will be more accurate and in at least two cases would have increased the Program's gross reported savings.

- **Add fields in the buildings portal to allow entry of more fuel types** - The SBC Operating Plan reports only for total non-electric MMBtu impacts, not by fuel type. There is a single field in NYSERDA's tracking system for aggregate MMBtu fossil fuel impact and a corresponding single descriptive field for the associated fossil fuel type. Multiple fuel types can only be recorded in a notes section of the database. Large studies and particularly CHP studies often impact multiple fossil fuel energy sources. Separate fields for all fuel types should be considered for the most accurate estimation of energy savings by fuel
- **Collect and retain electronic files -** Require that service providers deliver PDFs of completed studies and that they submit the Excel analysis files and building model input files (e.g., eQUEST .INP and .PD2 files) and archive them for at least 7 years. Keep electronic copies of the FlexTech study reports linked in the database for easy access. Early in the study start phase the study slowed down because electronic copies of the FlexTech reports were not available. Since the start of this evaluation, most of the sampled study reports have been scanned and connected to the appropriate study. However, significantly larger quantities of older FlexTech study reports still exist only in the hard copy format and a few were never found.
- **Incorporate premise identifiers –** This evaluation collected data to supplement program participation overlap analysis not in the scope of this research. The exercise would have been more effective and useful to later researchers and the data would be easier for them to use if each FlexTech study was associated with a unique premise identifier. A premise ID is a number that is associated with a physical customer address and remains constant over time regardless of occupancy. It is common to and transcends all programs and studies. Premise IDs are commonly used in the utility company environment and in fact it may be possible to collect and use the New York utility companies' premise IDs. Such an identifier also would aid in tracking for marketing and survey fatigue purposes.

Evaluation Recommendations

- **Select studies no older than 5 years in the next evaluation cycle** – The MAR survey results indicate that measure adoption rates plateau 5 years after delivering the study to the customer. Hence, in the next round of evaluations, the evaluators recommend using studies that are not older than 5 years beyond the survey date. As an example, if the next evaluation cycle were to be

conducted in 2012, then for this effort, the evaluators recommend picking up studies from 2007 onward. This will also help end use customers respond more effectively as they are likely to remember the details associated with the report.

- **Investigate and develop a more reliable method for the estimation of outside SO** - The outside SO rate is derived in this evaluation using the same method and survey questions as those in past evaluations. The final outside SO estimate ends up being based upon a small number of respondents (after dropping those that report no outside SO). It is a substantial estimate with uncertainty in many of its components. Further research is needed to develop a more reliable method that includes validity checks and is able to better estimate the full impact of the participating service providers on the market.

Section 1:

INTRODUCTION

The electric System Benefits Charge (SBC) is paid by customers of Central Hudson Gas and Electric Corporation, Consolidated Edison Company of New York, Inc., New York State Electric and Gas Corporation, National Grid, Orange and Rockland Utilities, and Rochester Gas and Electric Corporation. The New York State Energy Research and Development Authority (NYSERDA), a public benefit corporation established in 1975, began administering the SBC funds in 1998. The FlexTech Program is available to all electric distribution customers that pay into the SBC.

1.1 FLEXTECH PROGRAM BACKGROUND

The FlexTech Program funded through the SBC Operating Plans provides commercial, industrial, institutional, government, and not-for-profit sectors with objective and customized information to help customers make informed energy decisions. FlexTech's goal is to increase productivity and economic competitiveness of participating facilities by identifying and encouraging the implementation of cost-effective energy efficiency, peak-load curtailment, and combined heat & power (CHP) and renewable generation studies. Cost-sharing incentives are available to eligible participants for the following types of studies:

- General energy feasibility studies and technical support
- Peak-load reduction and load management
- Industrial and process efficiency analysis
- Data center efficiency analysis
- Energy procurement strategies
- Energy efficiency retro-commissioning
- Long-term energy and carbon management
- CHP & renewable generation study classifications
- Peak-load curtailment plans

The Program is designed to evaluate energy-related opportunities and the majority of a study must address electrical impacts. Program participants receive cost-shared analysis targeting their particular energy and business needs. Participants may use either their own independent consultants or FlexTech consultants under NYSERDA contract that are competitively selected to provide a statewide geographic distribution of needed technical services.

Smaller customers are eligible for walk-through energy audits, including a reimbursement of audit cost upon implementation of recommendations.

1.2 EVALUATION OBJECTIVES

The primary purpose of this FlexTech impact evaluation is to estimate the savings attributable to the Program. The evaluation needs to account for multiple layers of decision-making and performance in order to estimate the installed savings resulting from study funding, specifically: (1) what recommended measures study recipients implemented, (2) how much those implemented measures save compared to the study's predictions, (3) how NYSERDA's study funding influenced the recipients' decisions both to

receive the study and to implement measures, and (4) how participant experience influenced them to take additional actions beyond those recommended in the studies.

To achieve this goal of estimating attributable savings, the evaluation team determined:

1. **Measure adoption rate (MAR) -** A ratio that quantifies the percentage of study-recommended savings that customers chose to adopt. This factor solely addresses decision-making. It does not consider actual evaluated performance.
2. **Savings realization rate (SRR)** –A ratio that divides the evaluated savings for installed measures by the study-recommended savings for installed measures. The product of the MAR and SRR is the overall realization rate.
3. **Free ridership (FR)** –The likelihood the participant would have received the study or implemented the measure without the Program.
4. **Spillover (SO) -** The degree to which the customers' participation in the FlexTech Program influenced them to take additional actions to save energy.
5. **Outside participant spillover (SO)** – The degree to which vendors' participation in the Program influenced their practices with other non-Program customers.
6. **Nonparticipant spillover (SO)** – The degree to which the Program influenced non-participating end users and vendors to save energy through measures not funded by the Program.

The overall formula for net impact follows in Equation 1:

Equation 1: Net Impact Calculation

$$Net\ Impact = Gross\ Recommended\ and\ tracked\ savings * MAR * SRR * NTGR$$

where,

MAR = Measure adoption rate

SRR = Savings realization rate

NTGR (Net-to-gross ratio) = 1 – FR + SO

In the event that the study report's recommended savings differed from NYSERDA's tracking system entry of savings for the measure, the tracking system savings were used. This discrepancy rarely occurred. The research evaluated the impact of studies completed from January 1, 2007 through September 30, 2009. The MAR survey was designed to estimate the MAR for electric and natural gas measures at the program level within the 90/10 confidence/relative precision standard. The sample was a stratified random sample for the purposes of ensuring a representative sample; stratification variables were upstate and downstate regions, the year of the study, and the size of the study.[10] The on-site survey was selected using stratified ratio estimation to target a 90/10 confidence/relative precision level for the statewide program over the entire evaluation cycle. The net-to-gross (NTG) survey was a census attempt of decision-makers for adopting sites. Non-participant SO and other indirect program effects are within the purview of a separate study.

Some of the implemented measures received implementation funding from NYSERDA or other sources. Determining the amount of this programmatic overlap falls within the scope of a separate study and is not part of this impact evaluation. NYSERDA applies this overlap factor at the portfolio level.

[10] Downstate is defined as Con Edison's service territory. Upstate covers the service territories of all other investor-owned utility companies whose customers pay the System Benefit Charge.

As secondary objectives, however, this evaluation collected survey data regarding program funding overlap to aid later research as well as collecting information on the age of replaced equipment to help NYSERDA learn more about early replacement decisions.

1.3 REPORT FORMAT

The balance of this impact evaluation report is organized as follows:

- Section 3 provides details on the evaluation approach, including the sample design.
- Section 4 presents the evaluation results.
- Section 5 contains the conclusions and recommendations to program staff.
- Appendices A-L (listed below) follow section 5:
  - A: Participating Customer Measure Adoption Rate Telephone Questionnaire
  - B: Sample Design Details
  - C: On-Site Visit Checklist & Inspection Form Regarding Investment Practices
  - D: Participating Customer Net-to-Gross Telephone Questionnaire
  - E: TA Service Provider Net-to-Gross Telephone Questionnaire
  - F: Tracking System Description
  - G: Early Replacement Survey Instrument
  - H: Tabulated Measure Adoption Rates by Year and Category
  - I: Early Replacement Survey Data
  - J: Glossary of Acronyms and Definitions
  - K: MAR Round 2 Memo

Section 2:

EVALUATION METHODOLOGY

The section describes the methods used to estimate gross and net savings for the FlexTech Program. This discussion covers the overall approach, the sample design, the net-to-gross (NTG) analysis, and the calculation of total program savings.

2.1 OVERALL APPROACH

The evaluation team estimated three terms - measure adoption rate (MAR), savings realization rate (SRR), and net to gross (NTG) - to arrive at the final savings attributable to the FlexTech Program. The three terms were developed from data collected in four surveys:

1. A telephone survey of facility managers or engineers to determine the MAR and time of adoption for measures recommended in 303[11] studies completed between January 1, 2003 and September 30, 2009.
2. An on-site survey of forty-four participating facilities that had received studies between January 1, 2006 and September 30, 2009, completed the MAR interview and adopted at least one measure with measurement and verification (M&V) to derive the SRR for adopted measures.

   The product of these two factors – MAR and SRR – is the realization rate of study-recommended energy savings.
3. A telephone survey with a sample of decision-makers for participating customers to determine free ridership (FR) and participant spillover (SO). The respondent group was a subset of the original MAR survey respondents.
4. A telephone survey with a sample of their associated FlexTech service providers, also to help estimate FR and SO. The sample of studies selected for service provider calls was a subset of the decision-maker sample.

The MAR telephone survey was conducted from May through July of 2010. The subsequent on-site survey work began in December 2010 and was completed in March 2011. The attribution telephone surveys data collection period ran from March through May 2011.

The sample designs for the surveys used stratified random sampling. The sampling unit was the study total positive energy savings[12] on a source-equivalent basis.[13] Sufficient MAR data was

---

[11] Researchers completed telephone interviews associated with 301 Flextech studies. There were two additional studies for which the MAR results were already known, for a total of 303 completed MAR questionnaires.

[12] "Positive" energy savings means that only positive energy savings are included and additional energy use is excluded. Generation studies, for example, tend to result in large electric savings and also increased fuel use. This approach avoids including both negative and positive numbers in the calculations, which could have unintended consequences for the MAR.

[13] To convert to source MMBtu, the kWh savings for the electric measures were adjusted to account for savings at the source of generation. This approach avoids the potential pitfall of ending up with a sample that contains a disproportionate number of natural gas studies. The source factor provided by NYSERDA was 9,949.2 Btu/kWh, was based on a 3-year average (2006, 2007 and 2008), and includes a line loss factor of 7.2%. The number is based on natural gas only, since natural gas represents the fuel on the margin.

collected to report on adoption rates for energy efficiency versus cogeneration, electricity versus fossil fuel savings, and MAR by technology category.

This evaluation follows a conceptual framework similar to prior FlexTech evaluations in using the MAR, SRR, and NTG factors. However, greater funding allowed it larger samples, more stratification, greater precision, and enhanced rigor in assessing the realized savings and attribution of sampled studies.

2.1.1 Data Sources

Data was collected from five sources other than the survey interviews and site data collection. Those sources are discussed in brief below.

Program Data

NYSERDA provided access to the Program database and its dataset. Program data offered most of the detail required by the evaluation effort (e.g., the facility name, program type, measure level savings table, study cost, etc.), and missing data was obtained from the FlexTech studies and/or facility contacts during the MAR surveys. The recently completed studies in the database had PDF versions of the FlexTech reports, while the older studies had hard copies of the reports that had to be scanned and uploaded. While the database contained some of the correct site contact information, supplementary site contacts were identified by program and evaluation staff.

Previous Evaluation Data

Evaluators wanted to learn the long-term measure adoption rate (MAR). It was suspected that the rate did not plateau (no new measure implementation) for as many as 7 years after study completion. In order to determine long-term adoption rates, the evaluation team had to survey customers that received studies as long ago as 2003. The prior FlexTech evaluation surveys included some customers with studies all the way to the beginning of the Program (1998).

The evaluation team requested and received the 2005 and 2007 evaluation reports and the associated MAR and SRR calculation spreadsheets in order to leverage the prior study information and avoid contacting participants to ask them the same questions they answered previously.

The 2008 Impact Evaluation of the Largest Energy-Saving Projects comprised an in-depth, site-specific gross and net impact evaluation of each of the twenty to thirty studies with the highest savings in NYSERDA's C&I portfolio (Largest Savers). Since the sample frame for the on-site survey included studies initiated in 2006 and later, two of the Largest Saver sites were found to be included in the on-site and NTG samples for this evaluation. Data on adoption that had occurred since the prior evaluation was collected, but other results from that study—realization rates for implemented measures and responses to attribution questions—were leveraged and not verified again since that prior evaluation was highly detailed, customized, and fairly recent. Care has being taken to avoid double counting of savings and redundant application of evaluation ratios for these participants.

Billing Data

Billing data was required for a sample of the sites selected for on-site survey. Requests for this data were made directly to the site contacts and on a case-by-case basis were received.

Weather Data

Similar to the billing data, the Impact Evaluation Team reviewed individual study scope and decided which studies (site locations) required weather data for the selected engineering analysis. This data was obtained from NOAA and/or TMY3.

Analysis Files

In select instances, the Impact Evaluation Team acquired measure analysis files from study recipients or the FlexTech service providers.

2.1.2 Surveys

Four surveys were conducted for this evaluation:

1. MAR telephone survey
2. On-site survey of a subset of participants selected from the results of the MAR survey to estimate the SRR
3. Telephone survey of decision-makers for a subset of participants selected from the results of the MAR survey who installed measures, for estimating the NTG impacts
4. Telephone survey of matched participating FlexTech service providers for estimating the NTG impact, including those associated with the studies receiving the NTG telephone survey mentioned above and also a supplemental sample to reach desired sample size

ERS conducted the MAR and SRR surveys, while APPRISE[14] conducted the NTG surveys associated with the participant decision-makers and their matched FlexTech service providers. Megdal & Associates designed the NTG questionnaires and analyzed the results.

MAR Survey

The primary goal of the MAR telephone survey was to determine the MAR over time. The interviewer also attempted to collect data regarding whether the implemented measures received implementation funding from other NYSERDA programs or other sources.[15]

Researchers completed 301 telephone interviews in the late spring and summer of 2010 with recipients of studies that were completed between January 1, 2003 and September 30, 2009. There were two additional studies for which the MAR results were already known so the MAR portion of the survey was completed without a telephone interview, for a total of 303 completed questionnaires. The MAR reflects the percentage of savings from measures recommended in completed studies that were implemented in whole or in part. As part of this survey, evaluators tracked the funding of the implemented measures (NYSERDA, utility, federal, or other funding assistance) and also future implementation status of the measures. The response for the measure funding question was limited and hence is not presented in this report. Measures identified for near-term adoption were not counted in the MAR calculation.

If the contact reported that the measure was installed, never going to be installed, or was partially installed and not going to be further installed, the measure outcome was classified as "definitively resolved." Study recipients that had at least one measure that was not definitively resolved will

[14] APPRISE and their team of subcontractor's are contracted to provide telephone surveys for NYSERDA program evaluations.

[15] As is discussed later, the quality of response to this battery of questions was insufficient to use directly for overlap analysis.

continue to be contacted every 9 to 12 months to determine the status of the recommended and as-yet-uninstalled measures[16]. Appendix A presents the MAR telephone survey instrument.

The round 2 MAR telephone surveys were fielded a year after (May 2011) the round 1 MAR telephone surveys. The round 2 MAR sample focused on FlexTech studies completed in 2008 and 2009. Additional details about the round 2 MAR study can be found in Appendix K at the end of this report. The MAR estimates presented in this study are from the round 1 MAR telephone survey.

SRR On-Site Survey

The sample frame for the SRR on-site survey was developed from MAR respondents that adopted at least one measure for FlexTech studies completed from January1, 2006 to September 30, 2009. Site visits were conducted at a sample of forty-four participating facilities with installed measures to determine the SRR.

These site visits ranged in rigor from spot M&V to enhanced rigor M&V analysis. The rigor assignments were made based on the magnitude of the savings and the complexity of the measures involved. The goals of this survey are described below:

1. Field-verify the installation of measures claimed by the participant to be installed.
2. Determine and validate study or measure-specific baselines.
3. Estimate the savings for the sampled studies.
4. Compare these evaluation (ex post) savings estimates to the study estimated (ex ante) savings to derive the SRR.
5. Collect additional site-observed data as indicators of the decision-making process for that site (to be used as enhanced participant FR and SO).

NTG Survey

The Impact Evaluation Team estimated participant FR and SO for the FlexTech Program from surveys with participating customers and FlexTech service providers. This approach was based on an enhanced self-report process for both the largest FlexTech savers and those with the completed on-site surveys.

The NTG survey portion of the impact evaluation was designed to contact 121 participants who were a subset of the MAR survey respondents that installed at least one measure. In addition, all FlexTech service providers matched with the 121 participants were interviewed to complete the NTG analysis.

NYSERDA's evaluation survey contractor, APPRISE, administered the NTG survey and the goals of the survey are described below:

1. Provide study-specific FR and SO information.
2. Track the influence of other NYSERDA programs in assisting with the implementation of the recommended measures.

Although this survey is primarily designed to develop the NTG estimates, some responses assisted in understanding participant behavior and may therefore be useful to NYSERDA staff for program delivery purposes.

[16] The MAR round-2 calls were conducted in May 2011 and MAR round-3 calls are planned for December 2011.

2.2 SAMPLE DESIGN

Sampling was required for each of the three core components of the evaluation described above. Given that Flex Tech participants do not require follow-up from the Program to install measures, the MAR telephone survey was designed to determine the percentage of the recommended savings that were actually implemented. A nested approach to sampling was necessary and is described below.

- The initial (MAR) telephone survey was designed to determine whether measures were installed.
- The on-site and NTG surveys were conducted for sites found to have installed measures through the initial telephone survey.

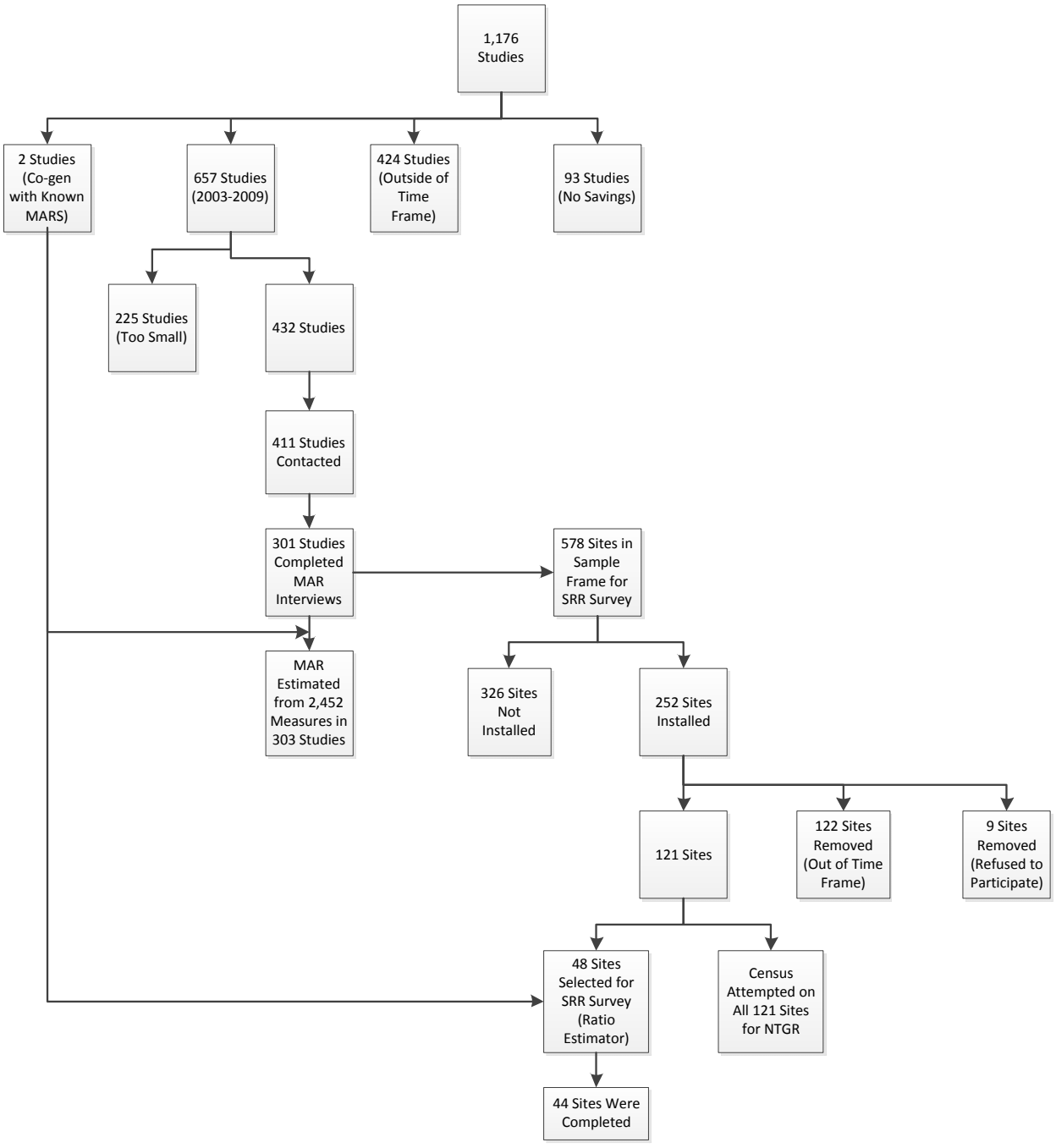
Thus, the results of the MAR telephone survey were used to develop the sample frame for both the on-site survey and the NTG telephone survey. Table 2-1 provides a summary of the overall sampling strategy for each factor to be estimated, including the data collection method, the population size, the sample size, and the way the results were applied to the program population.

Table 2-1. Summary of Overall Program Sampling Strategy

Factor	Data Collection Method	Population Size and Description	Sample Size	Application to Population
MAR (savings installed/ savings recommended)	Telephone survey of participants with FlexTech studies	Population: 1,176 sites Sample frame: 657 studies with FlexTech reports in 2003 and before October, 2009; of these studies, 225 were in the "too small to measure" stratum, leaving 432 studies	300 (411 of the 432 studies were contacted to obtain 300 completed surveys) Note – The evaluators ended up completing 301 surveys and including information for 2 large studies known to have been installed in the final tally. Information from 303 studies was included in the final tally.	Applied to all sites Stratified by age of report and upstate/downstate Estimated separately for electric, natural gas, and cogeneration savings
SRR (verified savings on site/ recommended savings installed for site)	On-site survey of participants (sites) with installed measures	Total sites with measures installed: 252 Sample frame: 121 sites with at least one recommended measure installed as verified through the MAR phone survey with FlexTech reports in 2006 or later; 2 large cogeneration studies known to be installed were also included. Removed from sample frame: 131 sites, consisting of 122 sites with measures installed with FlexTech reports prior to 2006 and 9 who refused on-site survey	50 sites (31 in census strata, 19 randomly selected) Note – The evaluators completed 44 on-site assessments which included the 2 large CHP studies.	Applied to all sites Stratified by upstate/downstate Estimated separately for electric and natural gas
NTG (free riders, participant spillover)	Telephone surveys of participants and the associated FlexTech vendors	Sample frame: all business owners and FlexTech vendors associated with the 121 sites with installations as verified through the MAR phone survey or other methods with FlexTech reports in 2006 or later	Participating business owners and FlexTech vendors associated with the 121 sites in the on-site sample frame (census) Note – completed 96 surveys.	Applied to all studies Census - application by strata to be determined

Figure 2-1 provides a graphical summary of the program sampling strategy.

Figure 2-1. Graphical Summary of Overall Program Sampling Strategy



2.2.1 MAR Sample

The first step in the sample design was to determine the method to be used for estimating the MAR. Stratified ratio estimation (SRE) was considered as a potential sampling method since it allows for efficient sampling design and generally requires a lower sample size for a targeted level of precision if there is a strong correlation between the tracking system savings and the implemented savings. However, a review of the previous evaluation done indicated that the correlation was not sufficiently strong to create an advantage for the SRE approach; hence stratified random sampling was chosen as the sampling method.

The MAR is calculated as the proportion of savings, based on self-reported installations of recommended measures in a Flex Tech study as determined through a telephone survey, divided by the savings associated with on-site verified installations. The precision/confidence target of 90/10 for the Flex Tech Program as a whole (statewide) was specified in the evaluation plan. The components of the MAR sampling plan are summarized in Table 2-2.

The sample was reviewed to verify that studies with savings from all major measure groups were represented in the sample. The MAR was calculated separately for natural gas, electric, and combined heat and power (CHP) studies.

The sample frame contained some studies that had been contacted for previous MAR evaluations completed in 2005 and 2007. All previously evaluated studies that fell into the census stratum were included in the sample and those in the random stratum were included in the sample to the extent that they were represented among the randomly selected studies.

Table 2-2. Summary of MAR Sample Design

Sampling Component	Sample Approach	Comments
Method	Stratified random sample	Results from previous evaluation suggest the correlation between adopted and claimed savings was too weak for stratified ratio estimation.
Primary sampling unit	Study	Some studies consist of multiple sites.
Definition of size	Maximum source MMBtu (electric or natural gas)	This approach allows studies with both electric and gas measures to be included in a single sample.
High level stratification variables	Upstate/downstate	90/10 precision/confidence target was set for each region; sample size was allocated proportionally.
Lower level stratification variables	Fuel type (gas only, electric/gas, electric only) and year of report	Stratification was designed to ensure proportional representation, not to achieve a specified precision/confidence target. Cut-offs were established by a review of the data.

A total of 657 studies were completed from January 2003 through December of 2009. The study size was determined according to the maximum source MMBtu (either electric or gas), and all recommended measures (both gas and electric) were reviewed for the selected studies. This approach allowed evaluators to analyze the study comprehensively and consider interactive effects between measures or groups of measures, as well as to include all sites with adoptions

from the MAR sample in the sample frame for the savings realization rate (SRR) site visits. More detail on the sample design is provided in Appendix B.

The sample size of 300 was estimated on the basis of professional judgment. This large sample was designed to be sufficient to allow stratification for the purposes of ensuring the sample is representative of the population and to estimate the MAR for the Program as a whole at or exceeding the 90/10 confidence/precision target.

Sample Frame Development

Since the MAR survey is intended to assess the installation rates over time, selecting the sample frame requires balancing between including older studies to enable assessment of how the adoption rates vary and restricting the survey to more recent studies to ensure that participants will be knowledgeable about the study and the implementation of recommendations. To address these competing objectives, studies completed from 2003 through 2009 were included in the sample frame. Studies met eligibility criteria for the MAR sampling frame based on the time passed since the participant received the Flex Tech report and the savings expected upon adoption of report recommendations. Specific reasons that studies were removed from the sample frame include the following:

- Studies that had no electric savings and no natural gas savings were excluded.
- Two large cogeneration studies were eliminated from the MAR telephone survey. These studies had been previously confirmed as installed and were verified through the SRR on-site survey.

Of the 657 eligible studies, 225 were found to comprise only 3% of the total recommended savings and these studies were excluded due to the additional costs of contacting them in relation to their total contribution to the program savings. Thus, the final sample frame consisted of 432 studies. Table 2-1 above summarizes development of the MAR sample frame in the first row.

The upper level of stratification was region (upstate/downstate), with the sample size of 300 allocated roughly in proportion to the savings within each region, as shown in Table 2-3. Downstate studies accounted for about 23% of the studies and 24% of the maximum source MMBtu savings.

Table 2-3. Studies and Total Savings by Region

Region	# of Studies	Maximum Source MMBtu Savings (Electric or Natural Gas)	% of Total Max Source MMBtu Savings	Sample Size
Upstate	508	7,653,056	76%	230
Downstate	149	2,352,028	24%	70
Total	657	10,005,084		300

Within each region, studies were divided into size categories. The stratification of the studies by size is described in Table 2-4, which shows the number of studies, the maximum source MMBtu savings, and the minimum and maximum source MMBtu savings by study for each stratum.

Table 2-4. Project Sites and Savings by Size Stratum for the Upstate Region

Stratum	Sampling Method	# of Studies	Sum of Maximum Source MMBtu Savings (Electric or Natural Gas)	% of Maximum Source MMBtu Savings	Minimum Source MMBtu	Maximum Source MMBtu	Target Sample Size
Upstate Region							
0	None	176	223,370	3%	26	2,948	0
1	Random	262	2,463,639	38%	2,985	24,805	160
2	Census	70	4,966,047	59%	25,244	620,908	70
Total		508	7,653,056		26	620,908	230
Downstate Region							
0	None	49	68,741	3%	263	2,717	0
1	Random	76	817,253	35%	2,841	24,963	46
2	Census	24	1,466,034	62%	26,000	209,393	24
Total		149	2,352,028		263	209,393	70

MAR Sample Disposition

Survey protocols required that each site in the sample be contacted a minimum of six times at different times of the day and days of the week. Attempts were made to contact almost all of the facilities in the sample frame, with only 21 of 432 studies left at the completion of the survey. The overall response rate for the MAR survey was 71%, as shown in Table 2-5.

Table 2-5. MAR Survey Disposition

Disposition	Random Stratum	Census Stratum	Totals
Total number of studies in sample frame			
Completed MAR survey	233	72	305
Could not reach	84	22	106
In progress/not completed	21	0	21
Response rate	69%	77%	71%

Response rates were reviewed by groups to assess whether specific segments of the population were disproportionately represented in the survey responses. This analysis considered the impacts of region, age of the study, fuel type, and size stratum (random or census). Based on these results, the case weights were adjusted for non-response by region (upstate and downstate), size stratum, and fuel type (studies with gas measures either in combination with electric or only gas measures v. studies with only electric measures).[17]

[17] Lohr, Sharon L. *Sampling: Design and Analysis*. Duxbury Press: Pacific Grove, California, 2009, pp. 265-272.

2.2.2 On-Site Sample

The results of the MAR telephone survey were used to develop the sample frame for the on-site survey. While the MAR survey was conducted at the study level, some studies included recommendations made for more than one building or site (particularly school districts, municipalities, and national chain stores). The on-site survey sampling was done at the site level to facilitate the on-site survey and eliminate the need for additional sampling within studies that are comprised of multiple sites. Through the review of the study reports and the information gathered through the MAR survey, the sites associated with each study and the installed savings at each site were identified to allow for sampling by site for the on-site survey.

While the MAR included FlexTech reports provided to participants from January 1, 2003 through September 30, 2009, the sample frame for the on-site survey was restricted to studies completed from January 1, 2006 through September 30, 2009.[18] This decision was made to ensure that the participant would be likely to recall the FlexTech report and be able to provide sufficient technical details regarding the report and the implementation of the recommended measures.

After the stratification had been completed and the size categories defined, the sample size for the randomly-selected sites was calculated as follows:

$$n = \left(1.645 \frac{er}{\sqrt{D}} \right)^2$$

where,

1.645 is the number of standard deviations of variation assumed at the 90% confidence level, i.e., that there is a 90% probability of the mean falling within this range

er is the error ratio

D is the desired relative precision

Sample sizes were estimated using an assumed error ratio of 0.60 and a relative precision of 0.10 for the overall program SRR estimate. The assumed value of the error ratio is in the range recommended in *The California Evaluation Framework*.[19] The allowable relative precision of the random strata was adjusted to reflect the savings in the census stratum, which were assumed to have no sampling precision since no sampling was conducted, and the finite population factor was incorporated into the calculation.

Table 2-6 shows the on-site sample design including sampling components, sample approach, and supporting comments.

[18] The approach presented in the evaluation plan was to solicit participants who received reports in 2007 or later for the on-site survey. However, the sample frame developed from the MAR survey results turned out to be smaller than anticipated and thus participants with reports issued in 2006 were added.

[19] TecMarket Works, et. al. *The California Evaluation Framework.* Prepared for the California Public Utilities Commission and the Project Advisory Group, June 2004, pp. 327-339 and 361-384.

Table 2-6. Summary of On-Site Sample Design

Sampling Component	Sample Approach	Comments
Method	Stratified ratio estimation	Previous experience indicated that the realization rates associated with installed measures tends to be highly correlated to the recommended savings stated in completed FlexTech reports.
Primary sampling unit	Site (building) with installed measures	Some studies consist of multiple sites. Measure installation location was confirmed through the MAR telephone survey.
Definition of size	Maximum source MMBtu (electric or natural gas)	This approach allows studies with both electric and gas measures to be included in a single sample. A census of the large studies was included in the sample.
High level stratification variables	Upstate/downstate	90/10 precision/confidence target was set for each region; sample size was allocated proportionally. Due to the small sample size in the downstate region, a census of the larger studies accounting for 95% of the savings was selected.
Lower level stratification variables	Size	Stratification was designed to ensure proportional representation, not to achieve a specified precision/confidence target. The cut-offs for the strata were determined according to the method described in *The California Evaluation Framework* (2004).
Time frame	Participants with reports received from Januray 1, 2006 through September 30, 2009	
Outliers	Two large generation studies included in the census stratum	Two very large generation studies were removed from the MAR sample frame since they were known to be installed and were considered to be outliers. Both of these studies were added to the SRR sample to verify the estimated savings in comparison to actual operating conditions.

As with the MAR survey, the upper level stratification variable was region and the sample sizes were determined based on a review of the site-specific information. In the downstate region, excluding six sites accounts for 4% of the realized maximum source MMBtu and conducting a census of the remaining eight sites was the most feasible approach.[20] Table 2-7 shows the number of sites and total savings by region.

[20] This percentage represents a higher percentage of the savings than the cut off for the upstate studies (1%) due to the few number of studies in the downstate sample frame and the decision to include all of the studies in the census strata.

Table 2-7. Sites and Total Savings by Region

Region	# of Sites[1]	Expected Maximum Source MMBtu Savings (Electric or natural gas)	% of Total Expected Max Source MMBtu Savings	Sample Size
Upstate	105	667,074	65%	40
Downstate	14	356,978	35%	8
Total	119	1,024,052		

[1] This table does not include the two large cogeneration studies.

The stratification of the studies by size is described in Table 2-8, which shows the number of studies, the maximum source MMBtu savings, and the minimum and maximum source of MMBtu savings by study for each stratum. The strata of very small studies, consisting of thirty-five studies that make up a total of 1% of the adopted maximum source MMBtu in the upstate region, were excluded from the sample because verifying these smaller studies would require resources but would not contribute to reducing the uncertainty of the SRR estimates[21]. For the small sites in both regions, the overall savings realization rate was assumed to be 1.00.

Table 2-8. Sites and Savings by Size Stratum for the On-Site Sample Frame

Stratum	Sampling Method	# of Sites[1]	Sum of Adopted Maximum Source MMBtu Savings (Electric or natural gas)	% of Adopted Maximum Source MMBtu Savings	Min Source MMBtu	Max Source MMBtu	Target Sample Size
Upstate Region							
0	None	35	6,619	1%	2	610	0
1	Random	49	155,387	23%	709	7,994	19
2	Census	16	211,010	32%	8,214	20,610	16
3	Census	5	294,058	44%	29,148	92,513	5
Total		105	667,074		2	92,513	40
Downstate Region							
0	None	6	13,961	4%	149	3,324	0
1	Census	5	61,144	17%	5,035	20,000	5
2	Census	2	72,480	20%	35,887	36,593	2
3	Census	1	209,393	59%	209,393	209,393	1
Total		14	356,978		149	209,393	8

[1] This table does not include the two large cogeneration studies.

[21] The six smallest studies that account for 4% for the downstate savings were removed. This percentage represents a higher percentage of the savings than the cut-off for the upstate studies (1%) due to the few number of studies in the downstate sample frame and the decision to include all of the studies in the census strata.

Six of the forty-eight facilities in the sample declined to participate in the on-site survey or were census studies corrected from non-zero to zero MARs during initial SRR investigation, effectively removing them from the SRR sample.[22] The two large CHP studies were included in the SRR census stratum. These two studies account for 40% of the total adopted annual MWh savings for those studies included in the MAR sample frame; thus verifying the accuracy of the savings was a critical component of the evaluation. The final analysis was conducted with the forty-four sites that received complete on-site verification. Appendix C shows the on-site survey instrument.

2.2.3 Net-to-Gross Sample

The sample frame for the NTG surveys was the same as the sample frame for the on-site survey, i.e., studies with at least one installed measure and FlexTech reports provided to the participant from 2006 to 2009. As discussed above, the sample frame consisted of 121 sites. The FlexTech evaluation plan specified that the M&V team would conduct NTG telephone surveys of 130 participating studies and the estimated seventy FlexTech engineering firms or vendors associated with these studies. These sample sizes were designed to meet the 90/10 confidence/precision target for the upstate and downstate regions. This approach ensured that the NTG surveys were targeted toward participants who followed at least one of the recommendations in the FlexTech report and also that the installations were recent enough to expect that the participants were able to recall receiving the FlexTech report and the process of implementing the recommendation(s). Given that the sample frame was smaller than the original sample size, a census of all 121 sites and associated FlexTech engineering firms were included in the NTG sample. While the number of completed surveys was lower than anticipated in the evaluation plan, attempting to contact all potential respondents in the sample frame eliminated the potential for sampling error. Appendices D and E present the participant owner and TA service provider survey instruments.

2.2.4 Sampling Conclusions

Estimating gross savings for the FlexTech Program has two components: the MAR and the SRR. Program savings were estimated as follows in Equation 2:

Equation 2. Verified Program Savings Calculation

$$\textit{Verified Program Savings} = \textit{MAR} * \textit{SRR} * \textit{Study Recommended Savings}$$

The MAR was calculated from the results of the telephone survey, which was designed to ascertain which measures were installed. Additional details on the MAR calculation methodology can be found in subsequent sections of this report.

The MAR is assumed to change over time, as older reports are more likely to have savings implemented than more recent ones. Therefore, the MAR sampling frame was stratified by the year of the report.

The results of the on-site survey were used to calculate the SRR. Additional SRR calculation details are provided in subsequent sections of this report. The sample frame for the on-site survey consisted of the sites with installed measures as determined through the MAR survey or other methods (e.g., the prior survey results) for FlexTech reports issued in 2006 or later. Since the purpose of the on-site survey was to estimate SRR, i.e., the relationship between the savings estimated in the FlexTech reports as compared to the actual savings acquired at the site, the

[22] This adjustment was necessary for sites when the participant responded to the MAR telephone call indicating that measures had been installed, but further investigation determined that this was not the case.

sample design was based on the assumption that the accuracy of the savings in the earlier reports is substantially the same as the later reports. Thus the SRR is expected to be the same over the entire period of program implementation and was applied to all studies.

The relative precision of the MAR estimates was calculated using the standard calculation, as given below in Equation 3.

Equation 3. Relative Precision Calculation

$$rp = \left(1.645 \ \frac{cv}{\sqrt{n}} \right)$$

where,

cv = The coefficient of variation

n = The sample size

The finite population correction factor was also incorporated into the estimation of relative precision. The method for estimating the precision of the SRR was similar, using the error ratio rather than the coefficient of variation. For both surveys, the weights were adjusted for non-response by using weighted-class adjustments.[23]

The precision of the total gross savings was adjusted to account for the two-stage sampling process for the MAR and on-site surveys. Since the sample frame for the on-site survey was a subset of the respondents from the MAR telephone survey, the worst case scenario was assumed, i.e., that the two sample frames are completely correlated. In this situation, the combined precision of the total gross savings is approximated by Equation 4 below.

Equation 4. Combined Precision of Total Gross Savings

$$P_{combined} = P_{MAR} + P_{SRR}$$

where,

$P_{combined}$ = The prccision of the total gross savings

P_{MAR} = The precision of the estimated MAR

P_{SRR}= The precision of the SRR

2.3 DATA TRACKING

The FlexTech Evaluation Team used a web-based tool called Salesforce to track the MAR telephone survey calls, document the survey results, and track the study's progress. Salesforce provides the means of tracking the progress of many events associated with multiple studies in a central database that is accessible from multiple locations. The developers originally built the software as a contact management tool for enterprise-level marketing departments, but it is a powerful tool for evaluation field management as well. Both sales and evaluation M&V tracking systems require monitoring of groups of people managing studies through milestones and recording contacts and their outcomes. Salesforce enabled our team to track the results of the contacts associated with more than 400 customers using a common tool that could be accessed by

[23] Lohr, Sharon L. *Sampling: Design and Analysis*. Duxbury Press: Pacific Grove, California. 1999, pp. 266 - 268.

the entire team. It tracked both evaluation progress and the implementation status of the recommended energy efficiency measures with the ability to store and edit contact information, measure-related data, and survey responses in real time. It is also scalable for the future.

Appendix F discusses this tracking tool in detail and how the evaluation team used it for this study. In this report, the Salesforce database is referred to as the "tracking system."

2.4 MEASURE ADOPTION RATE

2.4.1 MAR Evaluation Approach

The study-level MAR represents the percent of savings of measures recommended in completed studies that have been installed. The MAR was calculated as shown in Equation 5.

Equation 5: MAR Calculation

$$MAR_n = \frac{\text{kWh}_{Installed\_in\_a\_year\_n}}{\text{kWh}_{\text{Recommended}}}$$

where,

n = Years between study completion and installation

MAR = Measure adoption rate

$kWh_{Installed\_in\_a\_year\_n}$ = Total kWh savings estimated in the FlexTech studies for all customer-reported installations in year n

$kWh_{Recommended}$ = Total kWh savings for all measures recommended for implementation in the FlexTech studies

A MAR of 1 would indicate that customers had installed all recommended measures in their FlexTech studies.

The MAR estimate was tabulated in a matrix format as shown below:

Year Bin	0	1	2	3	n+
Measure adoption rate	MAR0	MAR1	MAR2	MAR3	MARn

where,

MAR0 = MAR for year 0, meaning the first 12 months following study completion

MAR1 = MAR for year 1

MAR2 = MAR for year 2, etc.

Year X is the difference between the year in which the study was delivered to the customer and the year in which it was actually installed. It is likely that a study delivered to the customer in 2008 will have some measures in year bin 0, some in year bin 1, and some in year bin 2 depending on when they were actually installed. This method captures the rate at which the measures are adopted over time.

Units of kWh are presented in the MAR equation above. In the case of non-electric (fuel) savings, the same equation was used, with MMBtus replacing the electric kWh. These results are also separable by upstate/downstate, energy efficiency, measure technology category, and on-site generation.

The final MAR for each reporting cycle was the total installed kWh savings for that entire reporting cycle divided by the total recommended kWh savings for that reporting cycle.

The MAR telephone survey followed standard survey methodological procedures:

1. An advance letter was sent via postal mail to the Facility Manager or other identified contact at each study site prior to conducting the telephone calls.
2. The interviewers called at least six times at different times of the day, different days of the week, and over multiple weeks.
3. In the case of studies that included more than fifteen recommended measures, prior to the interview, the interviewers grouped measures that represented over 80% of the savings, creating a manageable list of measures to investigate. According to the interviewers, less than ten measures per interview were deemed as appropriate for a 30-minute interview. In the case of studies that had a majority of its measures of equal or similar magnitude savings (all converted to MMBtus), a random sampling tool was used to select ten measures that would be queried during the telephone survey.

Several respondents chose to respond through email exchanges. Overall the response rate exceeded 70%, and interviewers used the tracking system to enter responses.

Due to the highly technical nature of the questions and dialog, engineers rather than social science research professionals conducted the interviews. Inspections and subsequent in-person interviews during the site visits revealed that a material proportion of the MARs reported by respondents in the telephone survey was incorrect. It is believed that the complexity of the questions caused the discrepancies and that the erroneous responses would have been greater without technical interviewers. Furthermore having the same engineer conduct the interview and visit the site enabled coherent interpretation and correction of responses when necessary. It also is possible that fewer of the discrepancies would have been discovered otherwise.

The final step in MAR calculation was to develop a correction factor based on the observed / telephone interview-based adoption rates in the SRR sample. The correction factor was calculated based on the weighted average percentage of source-based savings that was corrected in the SRR sample.

2.4.2 Comparison with Prior MAR Evaluation Approach

The prior evaluation efforts completed in 2005 and 2007 used the same Equation 5 presented above to calculate the MAR. The most significant difference between this evaluation's MAR calculation method and the prior method is that this evaluation's questionnaire asked respondents to estimate survey installation date. Analysts constructed adoption rate curves as a function of time lapse between the study completion date and the installation date. In the prior study the adoption rate curve was constructed based on the time span between study completion and the study interview.

This survey's interview script (see Appendix A) differed modestly in terms of the questions posted. Differences in the core adoption rate question phraseology are minor enough to allow a comparison of results. Other differences relate to the number of completed MAR telephone interviews and follow up.

- The 2004 evaluation efforts were based on 139 MAR telephone surveys.
- The 2006 evaluation effort was based on 131 MAR telephone surveys, which included 99 studies from the 2004 evaluation.
- The prior evaluation samples were selected for an 80/20 confidence/precision.

- The current evaluation efforts completed 301[24] MAR telephone surveys with 90/10 confidence/precision for presenting the separate results for upstate/downstate and by energy efficiency and on-site generation.
- The prior evaluation accounted for any MAR-related discrepancies found during the on-site visits in the SRR calculations instead of restating the MAR.

2.5 SAVINGS REALIZATION RATE

2.5.1 SRR Evaluation Approach

The SRR is the ratio of the field verified energy savings to the phone verified energy savings. The SRR represents the percent of study-estimated savings for implemented measures that the evaluation team estimates as being actually achieved based on the results of the on-site survey and analysis.

The SRR calculation for electric energy is shown below in Equation 6.

Equation 6: SRR Calculation

$$SRR = \frac{kWh_{\text{Field Verified}}}{kWh_{Phone\ verified}}$$

where,

SRR = Savings realization rate

$kWh_{Field\ Verified}$ = Field-verified kWh savings (by M&V contractor or CIPP consultant) for all installed measures from FlexTech study

$kWh_{Phone\ verified}$ = kWh savings estimated in the FlexTech study for all measures that customers reported as installed during the phone survey

Note that unlike the MAR, the SRR is not time dependent.

The first step was to determine whether any measures in the sample had previously been evaluated or if they had received extensive M&V through program-funded activities. Specifically, the evaluator checked the following:

1. Whether the measure(s) was installed through the Enhanced Commercial/Industrial Performance Program (ECIPP) Tier 3, which requires substantive post-retrofit M&V. If so, the evaluation team reviewed the EF/ECIPP M&V documents for the analysis rigor and any deviation observed at the site and either used the EF/ECIPP program estimated savings associated with that study or leveraged it and improved upon the EF/ECIPP M&V estimates by collecting more information.
2. Whether the measure(s) previously had been evaluated as part of the rigorous 2009 Large Savers evaluation. If so, the FlexTech evaluator checked with the applicant for any changes that had occurred since the prior evaluation. If not, the prior evaluation results were used for those measures. If there had been changes, the large savers analysis was used as the starting point for modified analysis. There were two such studies.

[24] Researchers completed telephone interviews associated with 301 Flextech studies. There were two additional studies for which the MAR results were already known, for a total of 303 completed MAR questionnaires.

3. Whether the measure(s) had been evaluated as part of the prior FlexTech evaluation. This check only applied to the few studies in this sample that were completed in early 2006. If a measure was reported to be implemented in the prior evaluations, its prior reported MAR value was used. The prior MAR telephone survey measures were only surveyed if they were reported as not implemented or partially implemented.

After checking for prior evaluation, the evaluation manager developed preliminary levels of rigor assignments for each remaining study in the SRR sample according to a defined procedure.[25] This procedure helped the team optimize the use of available evaluation funds to minimize overall uncertainty. After rigor assignment the manager assigned lead evaluation engineers to each study.

The lead evaluation engineer and team developed a unique M&V plan and budget for each study in the SRR sample. In every case the MAR interview engineer led or was part of the SRR evaluation team. The team engineers recruited and visited the site, completed an interview, inspected the measures (and possibly non-measure systems for the NTG consideration), took spot measurements, installed loggers and returned several weeks later to collect them if the assigned rigor level supported such work, performed the savings analysis, and completed a site-specific evaluation report. Each M&V plan and report was subject to technical QC and each report was subject to editorial QC prior to completion.

Once the SRR had been calculated for each stratum, the results were combined to determine the SRR for the entire program. The two large CHP studies that were considered to be outliers were incorporated into the final calculations as follows:

- The SRR was calculated for all studies in the on-site sample by stratum and then combined to determine the SRR for the on-site sample (a subset of the MAR survey respondents).
- The savings for all studies in the total MAR sample frame (without the outliers) were aggregated and the MAR was applied to determine the total adopted savings as estimated by the Program; then the SRR was applied to determine the total gross savings excluding the outliers.
- The recommended savings for the outliers were added to the total adopted savings for the MAR sample frame, the verified savings were added to the total verified savings for the MAR sample frame, and the program SRR was recalculated.

2.5.2 Comparison with Prior Evaluation Approach

The SRR equation used in the prior evaluations was the same as the current evaluation equation. The following are differences between the current and prior evaluations concerning the estimation of SRR:

1. The SRR for the 2005 evaluation was based on the results at fifty-nine sites. The prior evaluation team conducted visits at thirty-four of the fifty-nine sites and used the ECIPP program results for the remaining sites.
2. The 2007 evaluation did not involve any site visits. The 2007 evaluation adjusted the 2005 SRR estimate based on some corrections.

[25] *Rigor Assignment for the On-Site M&V from the Measure Adopters among the Flexible Technical Assistance Participants,* from Jon Maxwell and Satyen Moray, ERS; Kathryn Parlin, WHEC; and Lori Megdal, Megdal & Associates, LLC (M&A) to Judeen Byrne, NYSERDA Energy Analysis and the Evaluation Staff of the New York Department of Public Service. November 16, 2010.

3. The exact rigor of the M&V process used in the 2005 evaluation is not well understood.
4. The current evaluation sampled forty-four sites for on-site surveys to estimate the SRR. Of the total, nineteen were assessed at the basic and enhanced analysis rigor involving deployment of multiple data loggers and analyzing data. The remaining sites were assigned spot measurement rigor, which still involved physical inspection of the installed measures and taking spot readings.
5. The prior evaluation sampling for the on-site work was based on 80/20 precision while the current evaluation is based on 90/10 precision at the program level and for upstate/downstate.

2.6 NET-TO-GROSS EFFECTS

2.6.1 Methodology for Free Ridership and Participant Spillover

In addition to estimating gross savings, an equally important element of impact evaluation is to construct solid and defensible estimates of all impacts that are program-induced (rather than naturally occurring). This is often accomplished through assessing the proportion of savings associated with those who would have taken the actions without the Program (free riders) in comparison to program savings and the proportion of the savings from actions taken outside NYSERDA programs, but due to the Program (SO). The combination of these - the net-to-gross ratio (NTGR) - becomes the adjustment factor to derive net impacts.

The reliability for attribution relies more on construct validity than on sampling precision and this issue is the most important reliability issue for FR (as opposed to gross savings impact methods). The alternative of what would have occurred cannot be known with certainty. Survey inquiry is complicated as it is asking about conjecture of a theoretical alternative. Use of prior survey experience for specific question wording, measuring FR in more than one way, and obtaining market or other comparatives are several ways to increase the reliability of the estimate of the NTGR. Measuring FR in multiple ways can increase the construct validity of the estimate.

Consequently the primary self-report inquiries and FR algorithm for this evaluation replicate previous evaluations of NYSERDA SBC programs. The prior method was validated recently through one of the industry's most comprehensive in-depth quantitative and qualitative FR investigations: the 2008 Impact Evaluation of the Largest Energy-Savings Studies (Largest Savers).

In recent years, the energy efficiency program offerings in New York have expanded significantly with the addition of utility efficiency programs. Self-report FR methods and other FR methods will likely need to change in future evaluations to account for the range of efficiency programs being offered. Since this evaluation covers the program period of FlexTech study reports being provided 2007 to 2009, which predates the field implementation of the utility programs, replicating the previous SBC evaluation methods provides a defensible FR estimate for this evaluation and will establish a benchmark for comparison to future FlexTech evaluations.

The remainder of this section covers a general description of the method used to estimate FR and participant SO, the source of the estimate of non-participant SO, additional detail on the algorithms used to estimate FR and participant SO, and an explanation of the calculation of the combined NTGR.

2.6.2 Overall Method for Free Ridership and Participant Spillover

For program participants, assessing program-induced impacts involves producing two estimates:

1. **Free ridership (FR)** - The Program-supported measures (or the proportion of the savings) participants would have adopted within the same time frame in absence of the Program.
2. **Spillover (SO)** - Additional efficiency actions that are taken due to what participants learned or experienced through the Program. "Inside" SO occurs when energy saving actions are taken at the same study site, but are not done as part of the Program. "Outside" SO occurs when energy saving actions are taken at other sites that are not part of their program participation.

Participant FR and SO data were collected through telephone surveys with participating customer decision-makers and participating FlexTech Technical Assistance (TA) service providers.

The FR method for this evaluation improved upon the prior FR method by collecting additional data to create an enhanced self-report. This approach was employed to increase the construct validity of the FR estimate. The enhancements required gathering more information on the efficiency decision-making as described below.

1. **Additional decision-makers** - For large studies, extra effort was expended to investigate the involvement of other decision-makers, the level of their influence, and their contact information. The Impact Evaluation Team, through NYSERDA's evaluation survey contractor, attempted to survey all of these influential decision-makers. FR estimates, based upon the prior FR survey questions and algorithms, were derived for the primary customer contact and the influential decision-makers by study and combined based upon the influence of the different decision-makers and the type of decision-making process used in that firm.
2. **Consistency check against capital investment policy** - An on-site survey instrument was developed to be used for large studies as part of the on-site visits. This instrument was designed to gather information on the energy efficiency and actual capital investment policies employed by the firm based on observation of the systems found on-site. As part of this process, the Impact Evaluation Team engineers noted the condition of non-program systems, ranging from "duct-taped together" to "energy efficient or relatively new." This information was then compared to the responses to the FR questions in the telephone survey. (The on-site survey instrument is provided in Appendix C.)

This approach was designed to develop a better understanding of the decision-making process for energy efficiency improvements within these firms and triangulate FR estimates using additional sources of data.

The prior method was based on estimating the savings-weighted average free-rider factor individually for participants and service providers and combining these separate estimates using a savings weighted average. In the Largest Energy-Savings Studies evaluation, survey questions were designed to investigate the influence of the vendor and their role within the participant's decision-making process for each measure. This approach provided sufficient information to combine the customer and vendor responses on a case-by-case basis and develop the overall savings-weighted free-rider estimate. The program free-rider factor was also calculated using the prior method of averaging participant and service provider averages to compare the two methods. This analysis indicated that the difference in the overall free-rider estimate was quite significant. Based on the findings and recommendations of this previous evaluation, more detailed research

was conducted to differentiate between the free-rider estimate provided by the participating building owner and the estimate derived from the TA service provider survey for the study.

2.6.3 Free Ridership Prior Self-Report Algorithm

The process of estimating a participant's FR using the prior evaluation method has several components, as explained below.

- Direct FR is estimated by asking how the Program impacted the timing, likelihood, and share of measures installed, with likelihood and share asked by measure.
- Study-level FR is based on asking the participant to estimate, across all measures, the proportion of the total study savings that would have been achieved without the Program.
- A consistency check is conducted by comparing these preliminary FR estimates against other measures of program influence.

An initial FR estimate for each participating customer is determined from two different sets of questions. The first set, measuring direct FR, asks the respondent to assess the impact of the Program on the installation of individual energy efficiency measures. The second set of questions asks the participant to estimate, across all measures, the proportion of the total study savings that would have been achieved without the Program. This process provides the second FR estimate, which is then averaged with the direct FR estimate for the study to develop a preliminary average FR estimate for each participant.

To determine direct FR, participants are asked the likelihood of installation without the Program for each measure installed. Because the Program may be effective in increasing the quantity of efficient equipment installed, participants are also asked to estimate the share (proportion) of the efficient equipment that would have been installed without the Program, e.g., the percentage of the light fixtures that would have been upgraded to efficient fixtures if the participant had not received the FlexTech study. The responses to these two sets of questions are multiplied together to determine the FR rate by measure.

The third element of the direct FR analysis is timing. For the FlexTech Program, the key issue is *when* the participant would have conducted a similar study. The NTG survey included a series of questions to determine whether the participant was planning to take similar actions prior to enrolling in the Program, and these responses were used to establish the timing multiplier. This factor adjusts FR for participants who installed measures earlier due to participation in the Program. If the participant would have initiated the same type of study within 1 year, the direct measurement of FR described above is left unadjusted. If the participant responds that they would have eventually conducted a study but not for 5 years or more, direct FR is multiplied by zero, indicating that the participant is not a free rider. Proportional adjustments are made for responses between 1 and 5 years. The preliminary participant FR is calculated as the savings weighted average of the measure level FR estimates multiplied by the timing factor.

A consistency check was performed by comparing the preliminary FR estimates developed through the above process to the participant's responses regarding the influence of the Program.[26] This verification process is facilitated by the use of another set of questions that asked about the

[26] Over 20 years of experience in estimated self-report free ridership for energy efficiency program evaluation has set standards for quality FR measurement. One of these is to include additional inquiries and perform consistency checks across the inquiries. Estimating the hypothetical construct of FR based upon a decision that the participant might never have faced is quite difficult. The FR measure also needs to measure what would have occurred in the absence of the Program, not just what the participant "intended" to occur (as many good intentions do not actually occur).

influence of the Program on the overall study. The surveys contained three questions that are designed to measure the influence of the Program:

1. **Prior plans** - Participants are asked if they had any plans to install the high efficiency measures recommended by FlexTech. If so, they are then asked to describe any planning for energy efficient installations the firm had done prior to participating in the Program. The interviewer then assigns a score for this level of planning that ranges from 0 to 4 with 0 indicating no plans and 4 indicating that the high efficiency system was selected and budgeted.

2. **Program influence** - Participants are asked if the FlexTech Program influenced the type, number, or efficiency of measures installed. Those attributing influence to the Program are asked to describe the Program's influence on the decision to incorporate high efficiency at this site. The interviewer then rates the response from 0 to 4 with 0 indicating the Program had no influence and 4 indicating the Program was the primary reason that the high efficiency system was installed.[27]

3. **Program importance** - Participants are asked to rate the importance of the Program in their decision to install high efficiency measures. They are asked to rate importance from 0 to 4 with 0 indicating the Program was not at all important in the decision to install the high efficiency system and 4 indicating the Program was very important in that decision.

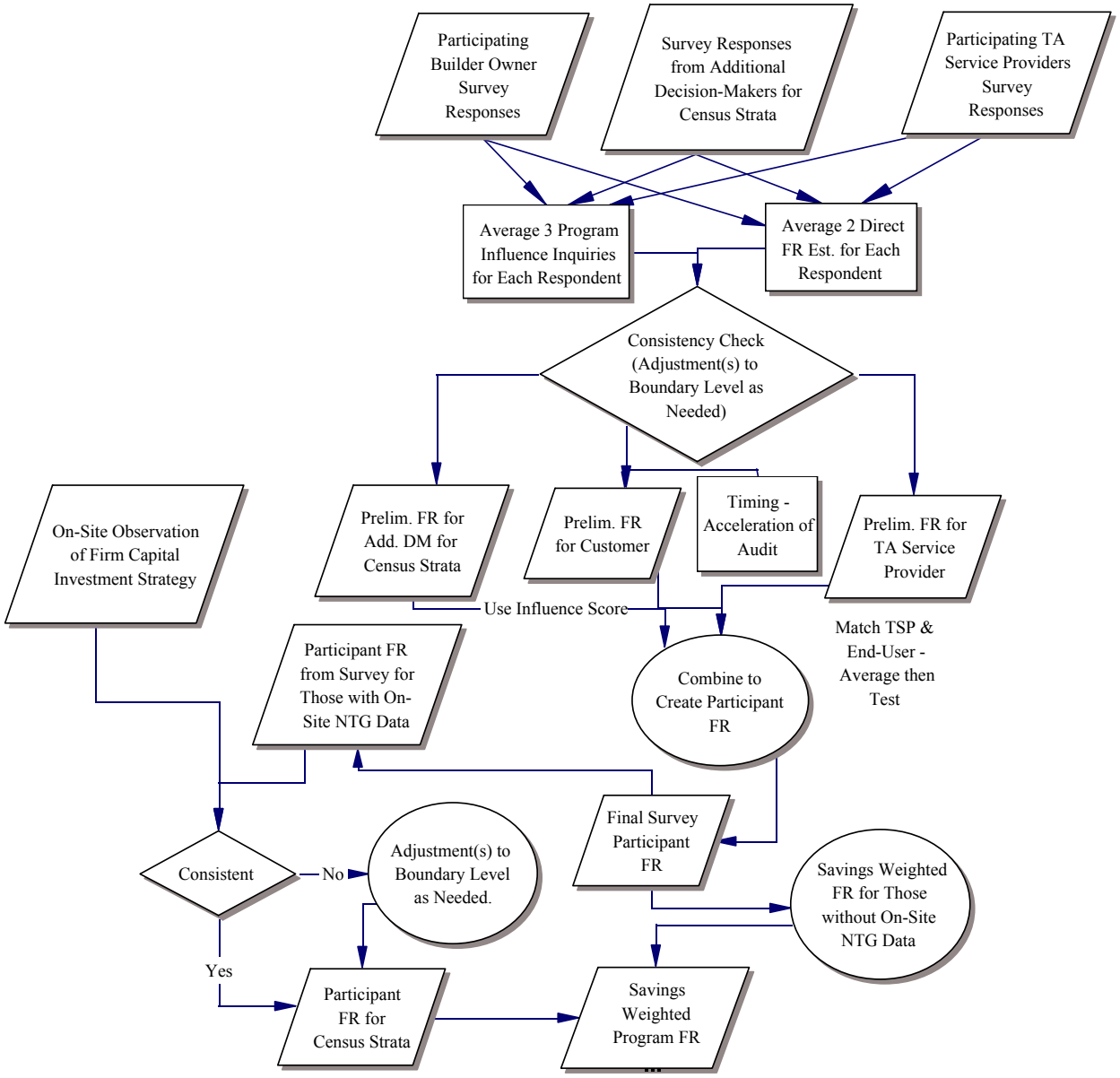
The responses from these three inquiries are averaged to determine the overall program influence score. This average is then converted into an upper and lower bound range of plausible FR values and compared to each participant's direct FR estimate discussed above. If the participant's direct FR estimate falls below the lower or above the upper bounds of FR based on the program influence questions, the final FR estimate is adjusted upward or downward to the edge of those bounds according to the influence score.

This algorithm for the calculation of a consistent participant FR estimate is used with the survey data from each respondent; it is the respondent's estimate of that participant's FR. The multiple respondents and additional data gathering are the elements that enhance this self-reported FR method from one where only the end user's telephone survey data is relied upon.

A schematic of the enhanced self-report approach used for this FlexTech evaluation is provided inFigure 2-2. It shows each of the components described above with what flow or combinations are used between them. (These are also described above.)

[27] A reference table defining each score was provided to the interviewer to enhance consistency among the raters (inter-rater reliability).

Figure 2-2. Schematic of the Enhanced Self-Report Components and How They Are Combined to Estimate Free Ridership for FlexTech



2.6.4 Participant Spillover Estimation Method

Spillover is defined as energy efficiency savings that are induced by the Program but had not directly been part of the Program. The building owner survey included questions to estimate participant inside and outside SO. (Participant inside spillover is when the Program induces the participants to take additional efficiency actions that were not part of the Program, i.e., the FlexTech study. Participant outside spillover is program-induced additional efficiency actions at other sites.). The participating service providers were asked about the Program's influence on their practice with other customers to estimate the participating TSP providers outside spillover.

Double counting is a concern when estimating SO from two separate sources (participating building owners and TA service providers). For example, it is possible that some of the building owners outside SO could be studies that are also outside SO for participating service providers

(the TA service provider for that building owner or another participating service provider). To the extent that this occurs, using both building owner and service provider estimates would overestimate SO. To ensure this does not occur, inside SO is estimated from the building owner surveys, and outside SO is based only on the service provider responses. The building owners' responses for outside SO are not used in deriving the Program's NTGR, but are reported as additional information.

The participant inside SO questions included the following[28]:

1. Were additional energy efficient measures installed at the same site due to the influence of the Program that did not receive direct program support?
2. If so, were the savings from the added measures higher or lower than the energy savings from measures that were supported by the Program? This question was followed by a request to estimate the savings from the additional measures as a percentage of the savings achieved through participation in the Program[29].
3. What is the percentage of these extra savings that could be attributed to the respondent's experience with the Program?

The outside study SO questions were similar, except that they were specific to efficiency measures installed at a different site.

The inside and outside SO questions were asked by fuel source (electricity versus natural gas) in order to maximize the accuracy that could be obtained in a telephone survey. The SO rates are calculated by combining the responses by fuel type, converting all savings to source MMBtu for electricity and natural gas MMBtu, and then estimating an overall value for both fuels. This dual approach was designed to improve the accuracy of responses while maintaining larger sample sizes for greater sampling precision.

The total SO rate was calculated by adding the inside and outside SO rates as seen in Equation 7.

Equation 7: SO Calculation

$$Total\ Participant\ Spillover = Inside\ Spillover + Outside\ Spillover$$

As SO occurs outside the bounds of the Program, it is not possible for the Program to track each occurrence. Conducting impact evaluations enables the identification and estimation of the SO impacts.

2.6.5 Non-Participant Spillover

The influence of NYSERDA's Flex Tech Program on the C/I sectors can easily overlap with the influence of 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 (NPSO) study applicable to C/I programs in 2005 and 2007[30], and an updated C/I NPSO

[28]Please refer to survey instrument Appendices E and F for the exact wording of the NTG questions.

[29] Efficiency program evaluation has found that many customers cannot identify their energy usage or the amount of energy saved accurately. The evaluation field has learned that it is more accurate to have respondents answer questions relative to their program energy savings. Then the analyst can use a relatively accurate estimate of those savings as the benchmark to calculate the energy savings relating to the survey inquiry.

[30] *Commercial and Industrial Market Effects Evaluation*, for New York State Energy Research and Development Authority, by Summit Blue Consulting and Quantec, October 2007.

study is currently underway. The C/I NPSO rate from this prior study will be used for all NYSERDA C/I programs until the results from the updated study are complete. Consequently, Flex Tech's NPSO rate used in calculating the NTGR is that from this prior NPSO study of 15%.

This evaluation did make inquiries concerning TA Service Providers expectations of NPSO to ascertain if they corroborated the occurrence of NPSO from the FlexTech Program. The TA service provider survey asks three questions to provide indicators of NPSO from the FlexTech Program.

2.6.6 Net-to-Gross Ratio

The above discussion describes two key components of constructing solid and defensible estimates of program-induced impacts, FR and SO, and how these are estimated in this impact evaluation. The surveys that gather the information needed, as described above, can be found in Appendix B.

The FR and SO rates are combined to produce a net-to-gross ratio that is applied to evaluation-estimated gross savings to produce net savings as seen in Equation 8.

Equation 8: NTGR Calculation

$$\textit{Net - to - Gross Ratio (NTGR)} = 1 - \textit{Free Ridership Factor} + \textit{Total Participant Spillover Factor} + \textit{C/I Non-participant spillover ratio}$$

This method enables the Program to account for participant net savings and SO savings, which ensures that the final net program savings estimate is all inclusive and credits the Program's impact in a precise and balanced manner.

2.6.7 NTG Survey Implementation

Three similar telephone surveys and a separate data collection effort within the M&V on-site survey were fielded. The NTG telephone surveys were a census attempt to survey participating owners and service providers for participants who adopted measures from studies completed in 2006 through 2009. (See Section 3.2 for more information about the sample frame and design.) The telephone surveys and on-site data collection are described below.

<u>NTG Telephone Surveys</u>

The three NTG telephone survey instruments were designed for the following groups:

1. Participating owners
2. Additional decision-makers in the participating firms that received studies with large savings
3. Participating TA service providers

The additional decision-makers for large savings studies were identified through the survey with the NTG participating owners, which asked respondents about the decision-making process relating to FlexTech participation and the contact information and influence level of other individuals in their firms who were instrumental in the funding and measure installation process.

The first two survey instruments were identical except that the survey of additional decision-makers did not contain the firmographics section, as that data had already been gathered during

the initial survey of participating owners. Advance letters to introduce the survey were sent to all potential primary participant contacts and the TA service providers before they were contacted.

Extensive checks were conducted by APPRISE prior to fielding to ensure that all skip patterns were correct and all question wording was comprehensible to respondents. Three participating owner surveys were conducted as a pretest to assess the wording of the survey instrument as well as to train interviewers at the call center; minor adjustments to the response options were made as a result of the pretests. Since the changes were minimal, one of the three pretests was included in the final data file. The other two cases were removed from our sample at a later date after being identified as sites that had participated in the 2008 Impact Evaluation Largest Energy-Saving Studies.

The interviews were conducted using a computer-assisted telephone interview (CATI) survey instrument. Two specially trained executive interviewers were used due to the complexity of the questionnaire and the need for the interviewer to exercise judgment in selecting questions and coding responses. These two interviewers conducted all of the surveys for this evaluation. Extensive training and quality control checks were conducted by APPRISE to ensure that the survey house had the necessary support to field the surveys.

The survey of participating owners was in the field from April 5, 2011 to May 6, 2011 and the calls to additional decision-makers were made from Monday, April 25, 2011 to Wednesday, May 25, 2011. The interviews with the TA service providers were conducted from April 26, 2011 to May 27, 2011. Interviewers called potential respondents during daytime weekday hours and calls were rotated between the morning and afternoon on different days of the week.

For the service provider survey, there were two types of respondents based on the number of FlexTech sites associated with the service provider:

1. TA service provider for one site
2. TA service provider for multiple sites

For TA service providers who conducted FlexTech studies at more than one site, survey administration procedures were slightly different after the initial survey was complete. The first time through, the respondent was asked all questions in the survey. For subsequent surveys, only questions pertaining to the unique FlexTech site were asked; these questions covered program influence, FR, and participant SO. General questions about the TA service provider were asked only once to reduce respondent burden, avoid redundant questions, and maximize the number of completed interviews.

APPRISE staff prepared and sent out letters to two difficult-to-reach TA service providers associated with multiple FlexTech sites. These two difficult-to-reach respondents conducted FlexTech studies for a total of eighteen sites in the sample frame (22% of the total sample frame). NYSERDA staff also reached out to these respondents via e-mail. Combined efforts resulted in one respondent completing the seven surveys for the sites where the firm conducted FlexTech studies.

A total of forty-eight interviews of participating owners and one pretest interview were completed and included in the final data file for these respondents, and forty-seven interviews of TA service providers were conducted. Table 2-9 shows the disposition of all sampled telephone numbers dialed for the participating owner and service provider surveys and provides the contact,

cooperation, and overall response rates. APPRISE provided survey implementation reports with the details of the survey implementation described in this section.[31]

The American Association of Public Opinion and Research's (AAPOR) definitions for sample disposition are used by APPRISE for the FlexTech NTG telephone surveys. (The definitions of the standard sample disposition rates are provided in Appendix J of this report.) The cooperation rate is the percentage of contact numbers that consented to an interview. The response rate is the number of completed interviews divided by all eligible sample units. The NTG Participating Owner survey achieved a cooperation rate of 93%, and a response rate of 64%. The NTG TA Service Providers survey achieved a 100% cooperation rate and an overall response rate of 69%. These rates and the detail counts are provided in Table 2-9.

Table 2-9. Sample Disposition for the FlexTech NTG Surveys of Participating Owner and TA Service Providers

Total Sample Used		**Number of Participating Owners**	**Percent of Participating Owners**	**Number of Participating TA Service Providers**	**Percent of Participating TA Service Providers**
		77	100%	82	100%
Excluded sample	Not working/unusable number	0	0.0%	2	2.5%
Not contacted	Respondent never available	6	7.8%	4	4.9%
	Answer machine	15	19.5%	16	19.5%
	Call back/left 800#	0	0.0%		
Unknown eligibility	No answer/busy	3	3.9%	1	1.2%
	Records not yet called/scr. not complete	0	0.0%		
Excluded business	Not eligible/not qualified	0	0.0%	12	14.6%
	Over quota	0	0.0%		
Refused/ break-off	Refused/ break-off	3	3.9%		
		1	1.3%		
Completed interview		49	63.6%	47	57.3%
Contact rate (53/74=.716)			71.6%		70.1%
Cooperation rate (49/53=.925)			92.5%		100%
Response rate (49/77=.636)			63.6%		69.3%

See the Glossary for definitions of Contact Rate, Cooperation rate and Response rate as defined by AAPOR.

The survey of additional decision-makers proved to be more difficult than the survey of participating owners. These decision-makers were identified by the original respondent of the participating owner survey but contact information for these additional decision-makers was seldom provided. Looking for contact information and re-contacting the original respondent failed to aid completion of these

[31]All of the disposition codes and rate formulae provided by APPRISE are consistent with the standards of the American Association for Public Opinion Research (AAPOR).The contact, cooperation, and response rates are the AAPOR #3 rates.

types of surveys. A census attempt was conducted for the NTG additional decision-maker sample. Three of these surveys were completed from a target of twenty-one.

On-Site Data Collection

The final survey used for the FlexTech NTG analysis was an observational component of the on-site M&V survey with some of the larger participants. The survey items were incorporated with the site visit checklist, and the data was collected by the engineer performing M&V on-site. The engineer asked whether the firm had a written energy efficiency policy and, if so, requested a copy. The engineer noted the presence and condition of energy-related equipment beyond those recommended in the FlexTech study. For these non-program measures, the engineer provided a condition and efficiency score from 0 to 5 with 0 representing equipment that was old and patched together and 5 representing high efficiency equipment in good operating condition. Further descriptions of these scores are shown in Table 2-10.

Table 2-10. Indicator of a Firm's Capital Investment Strategy from On-Site Observation

Condition/ Efficiency Score	Description/Guidance for Scoring
0	"Duct-taped" to make barely operational Very cost-effective to fix correctly or replace Unbelievable, looks like they are ready to go out of business
1	Old and obvious that replacement should have been undertaken to lower operating/repair costs
2	Operational but retrofit, for energy efficiency or repair costs, quite likely to be cost-effective.
3	Not high efficiency but not old and meets standard industry practice
4	High efficiency equipment as of a few years ago Payback for retrofit probably at least several years out
5	High efficiency equipment in good operating condition

2.6.8 Data Available for NTG Estimation

As discussed in Section 3.6.1, construct validity is the area of greatest concern regarding estimating free ridership. This issue is addressed through the use of multiple inquiries and consistency checks of a participant's responses. The series of direct self-report FR questions was designed to elicit explicit estimates of FR from the respondents. A participant's FR estimate was adjusted when the two direct FR measurements conflicted with questions regarding the level of program influence on their adoption decisions.

The enhanced self-report approach required additional data collection beyond the telephone surveys with the primary end user and TA service provider, including the following:

- Telephone surveys of additional influential decision-makers for the largest energy saving participants
- On-site data collection relating to the participating firm's capital and energy efficiency investment strategy
- On-site observations regarding the type and condition of the existing systems still in use, i.e., energy-related systems that were in place prior to the FlexTech report and in use at the time of the on-site survey

The NTG telephone surveys were a census attempt of the participating end users who adopted one or more of the recommendations from FlexTech reports issued in 2006 through 2009. A total

of forty-nine surveys were completed with participating owners, and the TA service providers were surveyed for forty-seven sites.[32] (Telephone survey implementation and survey disposition are discussed in detail in Section 3.6.7.)

The on-site NTG data regarding whether the firm had an energy efficiency policy was collected from forty-one sites and non-program systems was scored based upon its condition and energy efficiency at six sites. The availability of various types of FR information by number of sites is provided in Table 2-11 below.

Table 2-11. Number of Participating Sites with the Types of NTG Information Available for Them

NTG Data Source	Number of Sites with Usable Information[1]	Proportion of NTG Analysis Sites with Information from This Source[1] (Total Sites in NTG Analysis = 67)a
NTG participating owner telephone survey	47	70 %
NTG participating TA service provider survey	46b	69%
Additional decision-makers	1	2%
Information on whether the firm has policy on energy efficiency – from on-site	41	61%
Capital investment strategy as indicated by observations assessing non-program equipment during M&V on-site	6	9%

[1]It was part of the free ridership evaluation design for sites to have multiple sources of data.

[a] The 67 sites for the free ridership analysis consists of 26 sites that have both end user and TA Service Provider responses, 21 that only have end-user responses, and 20 with only TA Service Provider responses. Most of these sites include on-site data.

[b] Participating TA service providers completed surveys for one to seven adopting sites they had worked with. Survey data was obtained on forty-six sites and that came from surveys with twenty-seven TA service providers.

Forty-nine end users completed surveys relating to forty-seven sites.[33] Survey data was obtained on forty-six sites that came from surveys with twenty-seven TA service providers. Across all sites, there were a total of forty-one matched TA service provider and end-user respondents who completed NTG surveys for twenty-six sites, as summarized in Table 2-12.

[32] Two end-user and one TA service provider surveys were excluded due to delayed installation of recommended measures.

[33] In addition to the primary contact, two additional decision-makers were interviewed for one end-user site.

Table 2-12. Number of Respondents by Actor Group across Sites

Actor Group(s)	Number of Respondents (n = 76)a	Number of Sites (n = 67)
End user and matched service provider	41b	26c
All end users	49	47
All service providers	27	46

[a] This is the sum of all end user and all TA service provider respondents.

[b] Some service providers participated in studies at multiple sites, thus this value includes each service provider only once.

[c]Two respondents for each study, end user and their TA service provider.

2.7 NET IMPACT

The research evaluated the impact of studies completed from January 1, 2007 through September 30, 2009. As presented earlier in Equation 1, the overall formula for net impact is:

$$Net\ Impact\ =\ Gross\ Recommended\ and\ tracked\ savings\ *\ MAR\ *\ SRR * N\ T\ G\ R$$

where,

MAR = Measure adoption rate

SRR = Savings realization rate

NTGR (Net-to-gross ratio) = 1 – FR + SO

The evaluation team calculated single overall estimates of SRR, FR, and SO for the evaluated studies. Because measure adoption occurs over time, the net impact varies by year. The evaluation calculates overall net impact and net impact factors by year. Because the studies continue to cause implementation into the future, the evaluation reports evaluated impact that has already occurred as well as the impact from past studies that is likely to occur in the future.

2.8 EARLY REPLACEMENT

In every FlexTech study that evaluators examined, the study author(s) used either "retrofit" or "new construction/replace on failure" as the basis of estimating savings to define measure type and calculate savings over the measure's presumed lifetime.

"Retrofit" means that savings are based on the difference in efficiency between the newly installed high efficiency equipment and the old inefficient equipment that was replaced. This savings is presumed to accrue each year until the end of the measure life.

"New construction/replace-on-failure" studies use a different baseline. Instead of using the efficiency of the old inefficient, replaced equipment to estimate the baseline efficiency, analysts use the currently applicable energy code. If no relevant code applies, current standard practice for new construction is used to define the baseline condition. Because energy codes and standard practices are continually increasing efficiency levels, this latter type of baseline typically results in less savings than the retrofit baseline. As with retrofit applications, the savings are assumed to accrue each year until the end of the measure life.

In a dual baseline assessment of a measure's savings, the initial savings are based on the retrofit definition. This higher level of savings is assumed to accrue for a number of years until the old

equipment would have failed. Then the savings for the remaining years of the measure life savings are based on a different and typically lower new construction/replace-on-failure baseline.

The NY DPS has directed New York energy efficiency program administrators to report and evaluate savings using the dual baseline approach when appropriate to do so, in May 2011.[34] Because the work being evaluated was completed prior to this date and the evaluation research plan also was completed before May 2011, this report's findings do not account for the concept of dual baseline.

However, NYSERDA's evaluation staff anticipated such direction for future programs based on prior DPS guidance.[35] With this in mind, the Impact Evaluation Team added questions related to this concept and tabulated the responses for use by future researchers in developing information on early replacement frequency of occurrence and on appropriate assumptions regarding remaining life of replaced but working inefficient equipment.

The Impact Evaluation Team added questions to the NTG survey instrument to collect data on the age of replaced equipment and the reasons for replacement. The two key questions were:

1. To the best of your recollection how old was the equipment you replaced?
2. Which of the following best describes your decision to replace the equipment? (Choices included "It was working but not as efficiently as newer models," "It was working but old and would probably need to be replaced in next couple of years anyway," and "It was not working.")

These questions were asked for all of the replacement measures in the study sample, up to five measures per site. In total, 151 non-blank responses were collected. The measures were binned into twelve different categories (e.g., controls, lighting, building envelope, VFDs, etc.). Analysts tabulated and cross-tabulated the results, without statistical weighting, and presented the results and related observations about the responses. The survey instrument details are presented in Appendix G.

[34] State of New York Public Service Commission, *Appendix M - Guidelines for Early Replacement Conditions*, created May 5, 2011 as a supplement to the October 2010 DPS *Order Approving Consolidation and Revision of Technical Manuals*.

[35] State of New York Public Service Commission, Case 07-M-0548 – Proceeding on Motion of the Commission Regarding an Energy Efficiency Portfolio Standard, *Order Approving Consolidation and Revision of Technical Manuals,* October 18, 2010, p. 9.

Section 3:

RESULTS

This section presents the results of the analysis by the major sections of analysis: measure adoption rate (MAR), savings realization rate (SRR), and net to gross (NTG). Within each section the results include a comparison with the prior study results for context, to the extent the results are comparable.

3.1 MEASURE ADOPTION RATE

The MAR measures the proportion of study-recommended energy savings that customers choose to implement. It is a decision-based parameter, not an engineering one. SRR accounts for the actual performance of the implemented measures relative to the study's predicted savings.

3.1.1 Current Evaluation

MAR[36] data was collected by telephone interview from a stratified random sample of 2003 through September, 30 2009 FlexTech completed studies.

Analysts aggregated the MAR results in two distinctly different ways. The first method of analysis estimates impact as a function of time elapsed since study completion. The second method of aggregation uses the same data but combines it differently, aggregating results using the sampling units and stratification basis in the sample design. The unit of classification regarding time is the study completion year. The result of the second method is an estimate of the MAR at the time of evaluation.

The first method uses the period of time for measures to be adopted. Program data tracked the study completion dates, and interviewers collected measure installation data for each of the 2,452[37] measures for which customers described the implementation status. Thus the elapsed time between study completion measure implementation is known for each measure. Using this information, the MAR as a function of elapsed time since study completion was calculated.

This information is powerful because it enables analysts to review the results over a long period of time and to combine the results from multiple study years into a single result, which tempers any boom or recession effects that may influence implementations rate associated with particular calendar years. Analysts used the sample design's expansion weight associated with the study multiplied by the source equivalent energy savings to represent the relative influence of each measure on the results.

The site visits for M&V revealed that the telephone-based MAR responses were not always accurate. Engineers identified forty-seven incorrectly reported measure statuses out of 151 measures in the SRR sample. About half of the incorrectly reported MARs were binary, meaning a measure declared to be installed during the phone interview (1.00 MAR) was not installed at all (0.00 MAR) or in one case, the reverse. The other half required adjustments to the percentages of the measures that were installed. Four adjustments increased the MAR, forty-three reduced it. After weighting and combining natural gas and electricity savings measures using the common

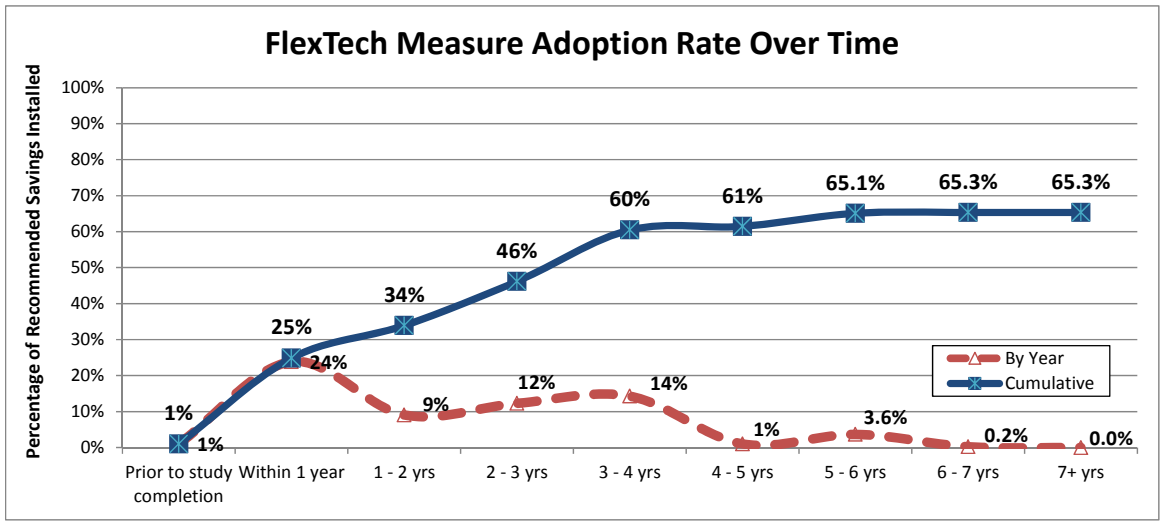
[36] Two rounds of MAR telephone surveys were conducted during the span of this evaluation. The MAR results presented in this report are from the more extensive round 1MAR telephone survey. The round 2 MAR telephone surveys were fielded a year after (May 2011) the round 1 MAR telephone surveys to update the implementation status on measures that were unresolved during the round 1 MAR telephone survey. Due to the timing of the round 2 MAR telephone survey, its results could not be incorporated into this study. The round 2 MAR telephone survey resulted in a very minor revision to the round 1 MAR. Additional details about the round 2 MAR telephone survey can be found in Appendix M at the end of this report.

[37] The 303 subject studies recommended a total of 2,519 measures. For 67 of the measures the respondent did not know the status, resulting in a total of 2,452 measures with known outcomes. The 67 measures were not included in the overall MAR calculations.

source-based Btu factor, the net overall site-based MAR adjustment factor was downward, (0.07). The telephone-based MAR results were multiplied by 0.93 to account for this correction.

Figure 3-1 shows the program overall adoption rate as a function of time elapsed since study completion.[38] The dashed line is the percentage of recommended savings adopted each year after study completion; the solid line depicts the cumulative percentage adopted.

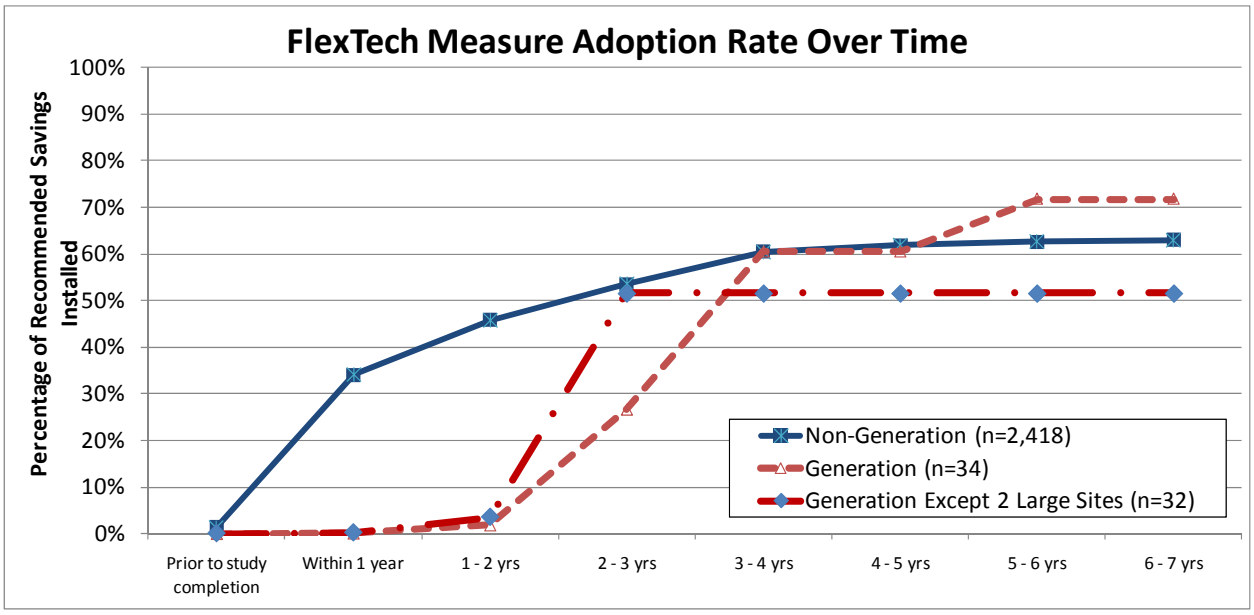
Figure 3-1. Program MAR Over Time



About 38% of the adopted measures are installed within one year of study completion, the mean time to adopt was 1.5 years, and 70% of measures are installed within 3 years. Measures continue to be adopted even in the 6th year after study completion before plateauing. The exception to this is CHP, as Figure 3-2 illustrates. It plots cumulative implementation rate and shows that most CHP studies require 2 to 3 years to implement and that two large CHP projects, required 3 and 6 years to complete. The numbers in the legend indicate the number of unique measures upon which the MARs were calculated.

[38] All numerical results associated with this and subsequent similar graphs are included in Appendix H.

Figure 3-2. Program MAR Over Time, Separately for Generation and Non-Generation Technologies



The FlexTech tracking system identifies the technology of each measure, a rich source of identifying information. Figure 3-3 further disaggregates the measure adoption rates by the technology for non-generation measures using this information (sixteen were unidentified). It shows that controls savings are by far the most frequently adopted measure type by study recipients - over 25% more controls are adopted than the next most readily adopted technology: lighting. Furthermore, study authors recommended controls measures 25% of the time for non-generation measures, which is more than any other technology. This controls trend is an interesting finding, as lighting often is perceived as the most common and readily adopted opportunity due to low uncertainty and lack of complexity. The pattern can likely be explained by the relatively low cost and fast payback time of many controls measures, but it also indicates customer willingness to implement measures that tend to be more complex. Envelope measures are the least adopted.

Figure 3-3. Program MAR Over Time, Separately for Type of Non-Generation Technologies

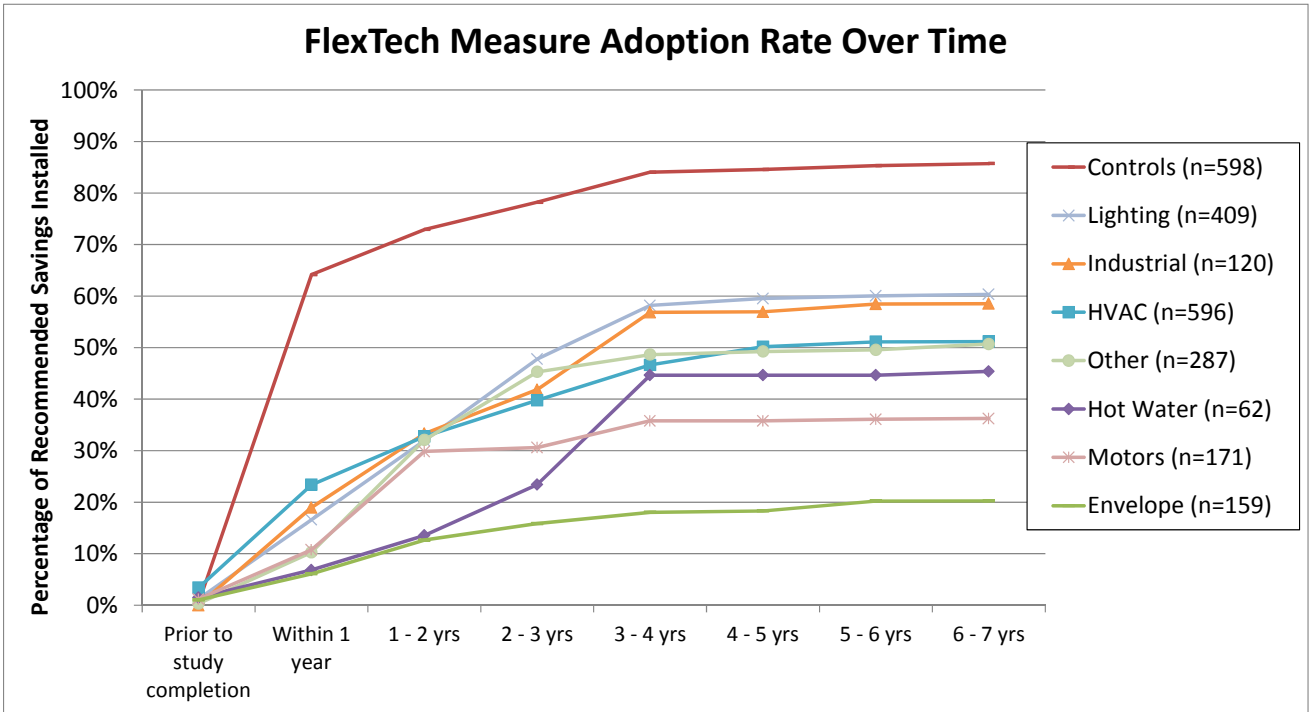


Figure 3-4 shows the adoption rates for upstate and downstate studies. The adoption rates differ significantly in the first few years but gradually converge before another modest departure.

Figure 3-4. Program MAR Over Time, Separately for Downstate and Upstate

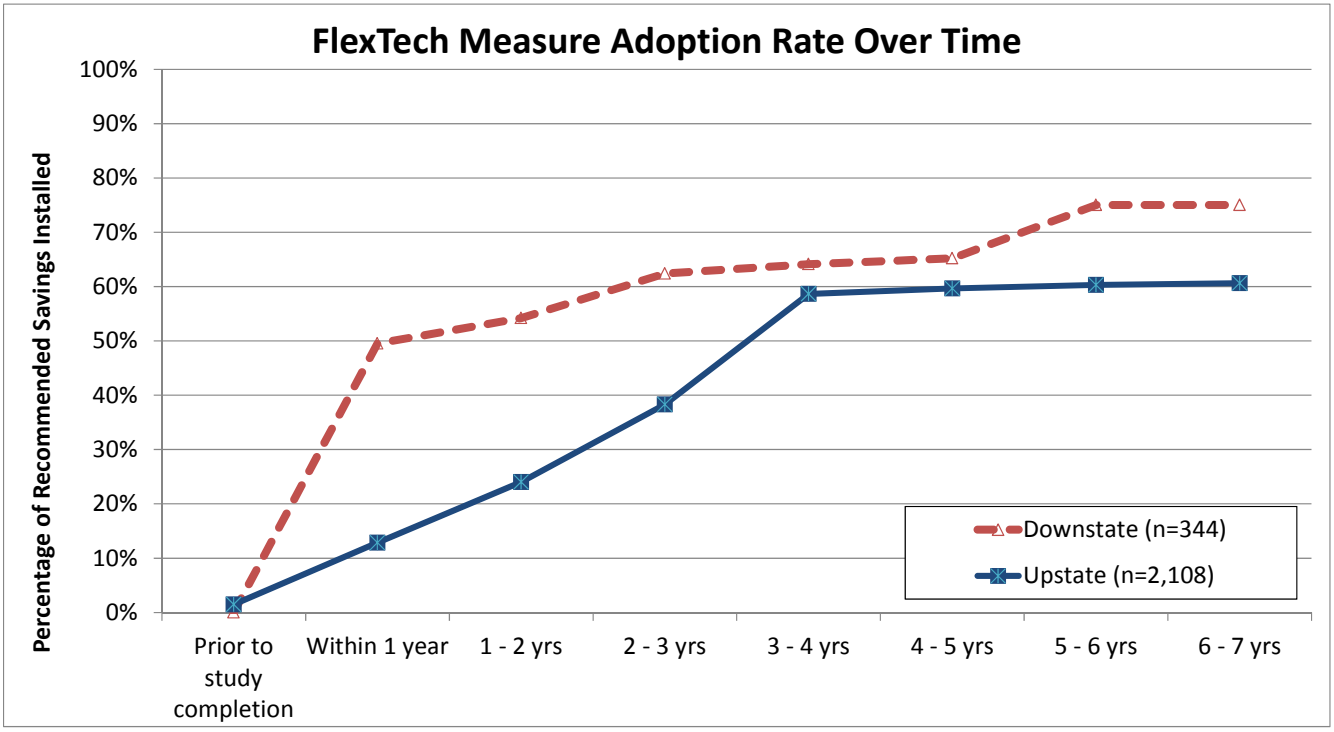
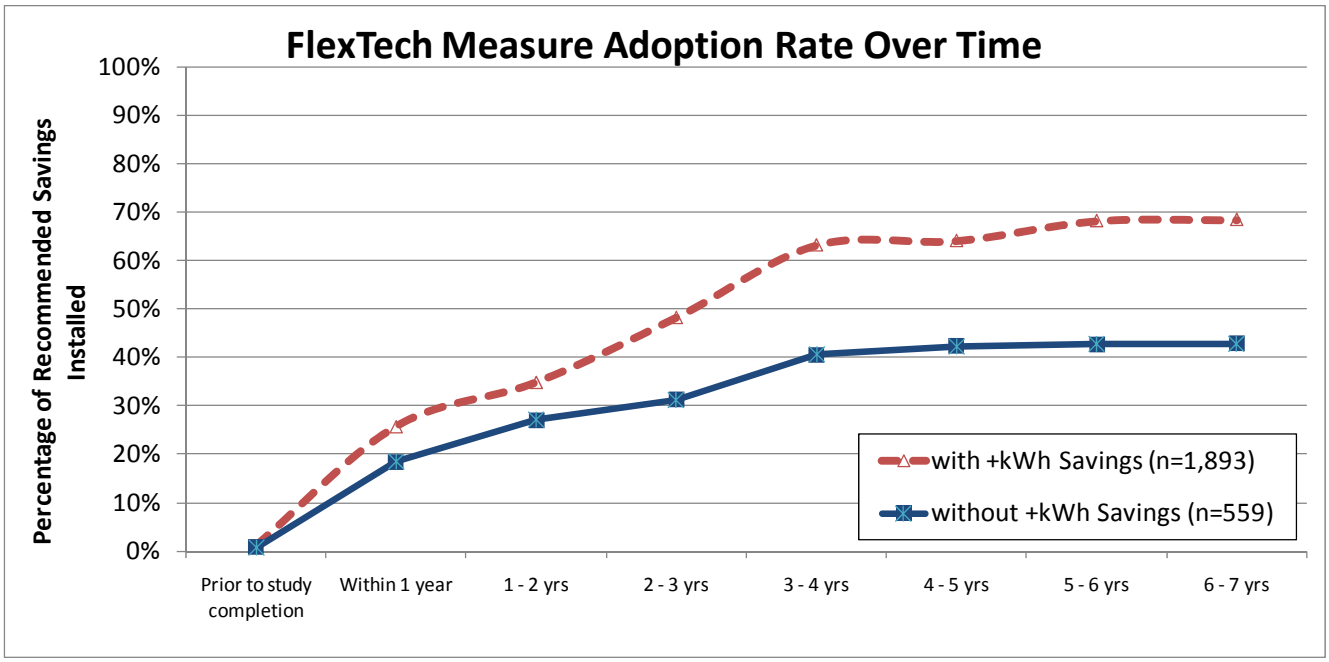


Figure 3-5 shows the adoption rates separately for measures that have electric energy savings and those that do not. For tabulation purposes, any measure that reported electric savings was classified as such; the other 559 measures saved only fossil fuel energy, according to study authors. There is substantial divergence in the ultimate MAR for these two categories. Historically gas efficiency measures have received less attention and incentive funding but this is gradually changing in New York and elsewhere.

Figure 3-5. Program MAR Over Time, Separately for Measures with and without Electricity Savings



Lastly, the evaluation team tabulated measure adoption rate over time solely with respect to the number of measures adopted and without regard for energy savings. Figure 3-6 summarizes the results with percentage of recommended measures at least partially adopted on the y-axis instead of percentage savings. Compared to Figure 3-5, the curve is smoother and plateaus at a lower rate. The lower plateau means that on average, customers adopted more of the larger savings measures and fewer of the lesser savings measures.

Figure 3-6. Program Number of Measures Adopted Over Time

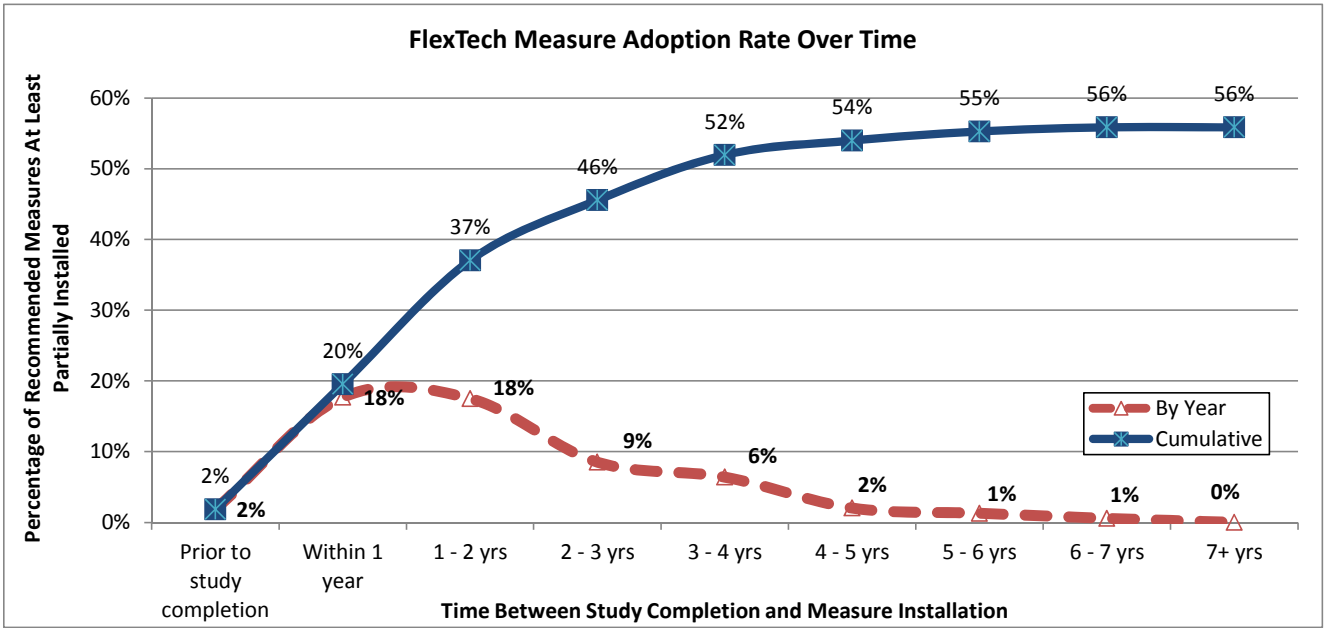


Table 3-1 summarizes the projected long-term MARs from the prior tables.

Table 3-1. FlexTech Projected Long-Term Measure Adoption Rates

Measure Type	Long-Term Projected Measure Adoption Rate
Electric energy efficiency	0.68
Non-electric energy efficiency	0.43
On-site generation	0.72
Overall	0.65

This first method of analysis estimates impact as a function of time elapsed since study completion. The second method of aggregation uses the same data but combines it differently, aggregating results using the sampling units and stratification basis in the sample design. The unit of classification regarding time is the study completion year. The result of this analysis is an estimate of the MAR at the time of evaluation.

This second method of analysis is important for two reasons:

1. It allows the evaluation team to report on the confidence and relative precision of the responses and variance of the data because, unlike the parameter time from study completion to implementation, the study year is a parameter that was identifiable in advance of the research and a basis of sampling. As Table 3-2 demonstrates, the results have little sampling error, less than 4% relative precision at the 90% confidence level for each of upstate and downstate on electric energy measures, downstate natural gas, and statewide on-site generation. The relative precision for natural gas upstate and consequently the total exceeds 10%, predominantly because of low MAR. The table also shows the population and sample counts.
2. It presents results in a format that can be compared to the prior evaluation because it is can be tabulate to present MAR as a function of study year and upstate/downstate, as is shown in the next section.

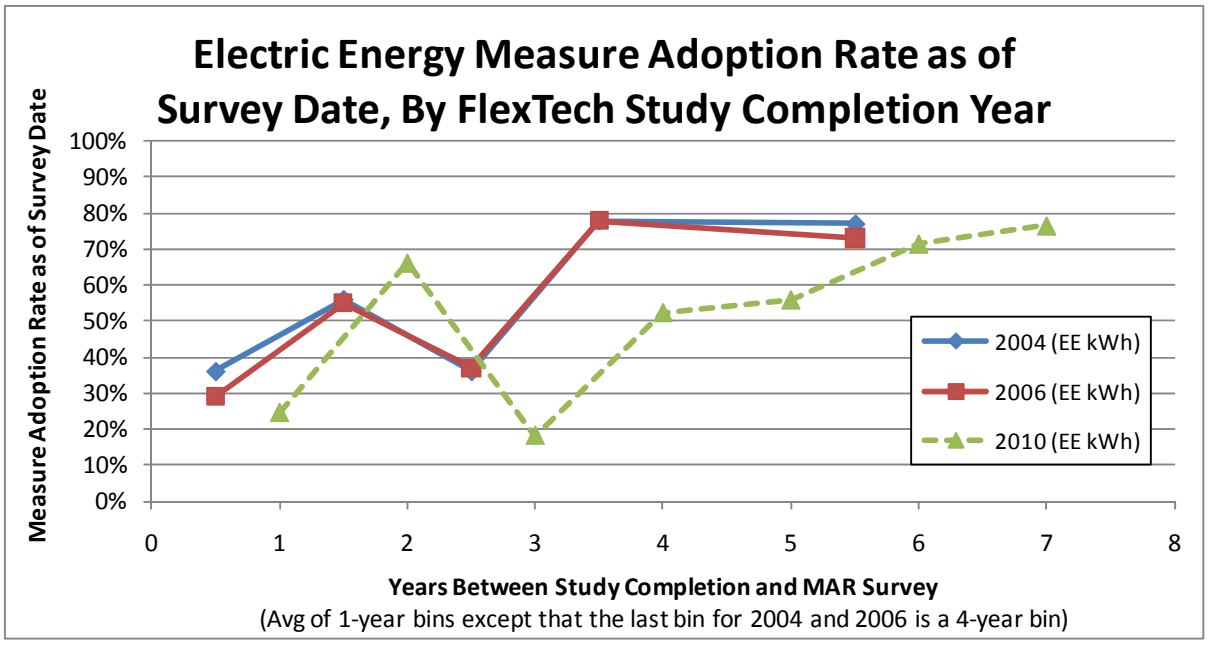
Table 3-2. Measure Adoption Rates at Time of Survey and Sampling Statistics

Parameter	Electric Energy			Natural Gas			On-Site Generation Electricity Only	Total Excluding Natural Gas Generation
	Upstate	Downstate	Total	Upstate	Downstate	Total		
MAR as of summer 2010	0.64	0.60	0.63	0.26	0.60	0.31	0.72	0.56
Number of studies in frame	506	149	655	298	60	358	41	655
Total sample	236	67	303	141	25	166	31	303
Standard error	0.017	0.026	0.014	0.029	0.004	0.025	0.015	0.014
Relative precision at 90% confidence	4.07%	6.64%	3.48%	17.58%	1.03%	12.44%	3.15%	3.73%
Coefficient of variation	0.380	0.330	0.369	1.269	0.031	0.975	0.106	0.485

3.1.2 Comparison with Prior Evaluation

Figure 3-7compares the results of this study with that of the two prior FlexTech evaluations for measures that save electric energy. All the evaluations analyzed fossil fuels and on-site generation but prior evaluations reported a single overall MAR rather than in year bins due to limited data.

Figure 3-7. Electric Energy MAR as of the Survey Date, for Prior and Current Evaluations



The patterns for the two studies are similar: They gradually increase over time but with a marked dip during 1 year. The 1/2-year offset of the curves is likely due to differences in binning rather than results. The three curves all plateau at similar adoption rates.

Error! Reference source not found. summarizes and compares the aggregate MAR to date for the current and prior FlexTech evaluations. To make the most similar comparison possible the table shows this study's results first without the site correction factor described in Section 4.1.1, as the prior study did not make this adjustment, and then the result with the factor.

Table 3-3. Overall Measure Adoption Rates as of the Evaluation Survey Date

Study and Parameter	MAR as of Evaluation Survey Date		
	Efficiency Measures Electric	Efficiency Measures Natural Gas	On-Site Generation
Prior evaluation – 2005	0.54	0.58	0.37
Prior evaluation – 2007	0.58	0.54	0.52
This evaluation – 2010 without the on-site correction factor	0.68	0.33	0.61
This evaluation – 2010 with the on-site correction factor	0.63	0.31	0.57

The comparison of results between this study and the prior one shows similar patterns of adoption over time. After minimizing methodological differences to the extent possible, this study found a higher electric MAR and lower gas MAR. The generation MAR is 10% higher. It is possible that the weighted average period between the study completion and the survey differed for the two evaluations, which would affect the comparison.

3.2 SAVINGS REALIZATION RATE

MAR measures the percentage of recommended savings that study recipients implemented. This parameter does not account for the actual performance of the measure. A 100% implemented measure may save more or less than the FlexTech study authors projected. Savings realization rate measures this variance. It is defined as the evaluated savings divided by the study's MAR telephone survey verified implemented savings (see Equation 9).

Equation 9: SRR Calculation

$$SRR = Savings\ based\ on\ evaluator\ M\&V\ /\ (Study - Predicted\ savings * M A R)$$

3.2.1 Site-Specific Energy Savings and Realization Rates

Engineers performed site-specific M&V on a sample of studies as described in the methodology section. Figure 3-8 tabulates facility types to help readers envision visited sites. Industrial participants were the most common facility type. Institutional customers (local government combined with education) were the biggest sector. Commercial facilities were least common.

Figure 3-8. SRR Facility Type Summary

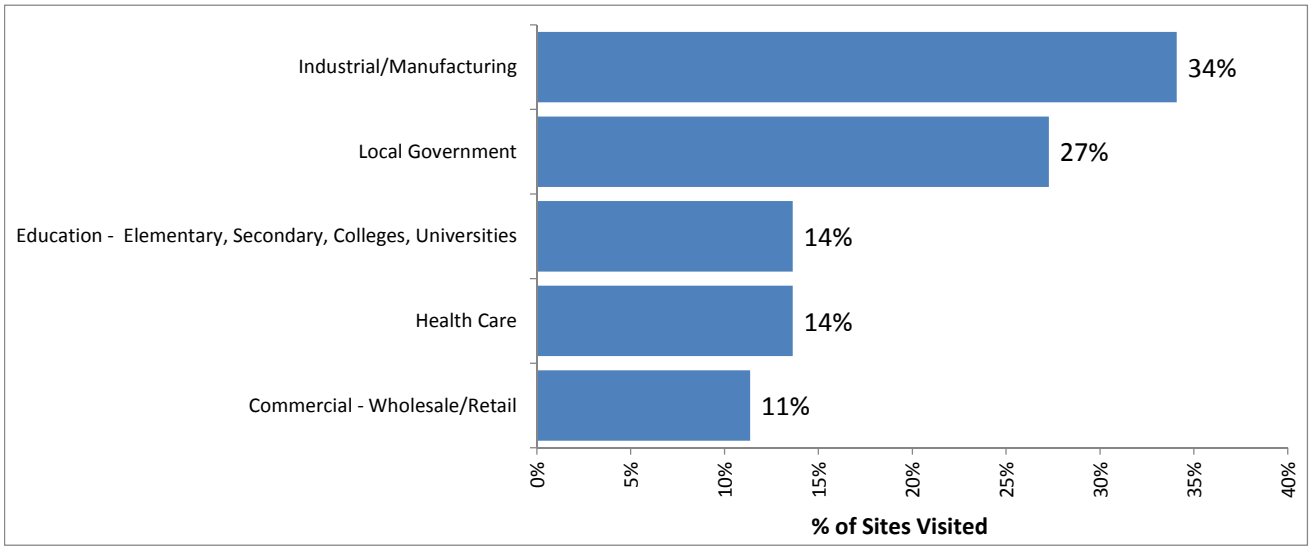


Figure 3-9 and Figure 3-10 illustrate the evaluated savings compared with that predicted by the FlexTech studies for electricity energy savings. Ideally, the evaluated savings would always match the study's recommended savings for implemented measures. This ideal is shown as a solid black line on the charts. Actual findings are plotted as points on the graphs. A pattern of points below the ideal line suggests a realization rate of less than one; points above the line suggest a realization rate greater than one. The error ratio measures the amount of scatter in the point distribution. If impact noted in the written FlexTech study differed from the savings in the tracking system, the x-axis of these charts and the realization rate analysis are based on the tracking system values.

Figure 3-9. Recommended and Evaluated Electric Energy Savings, All SRR Sites

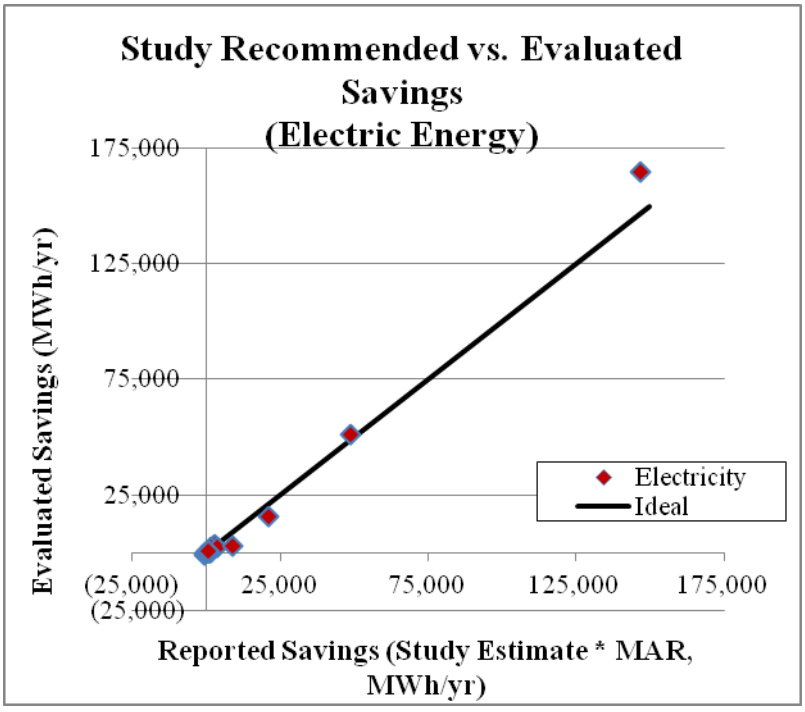


Figure 3-10. Recommended and Evaluated Electric Energy Savings, SRR Sites Excluding the Four Largest

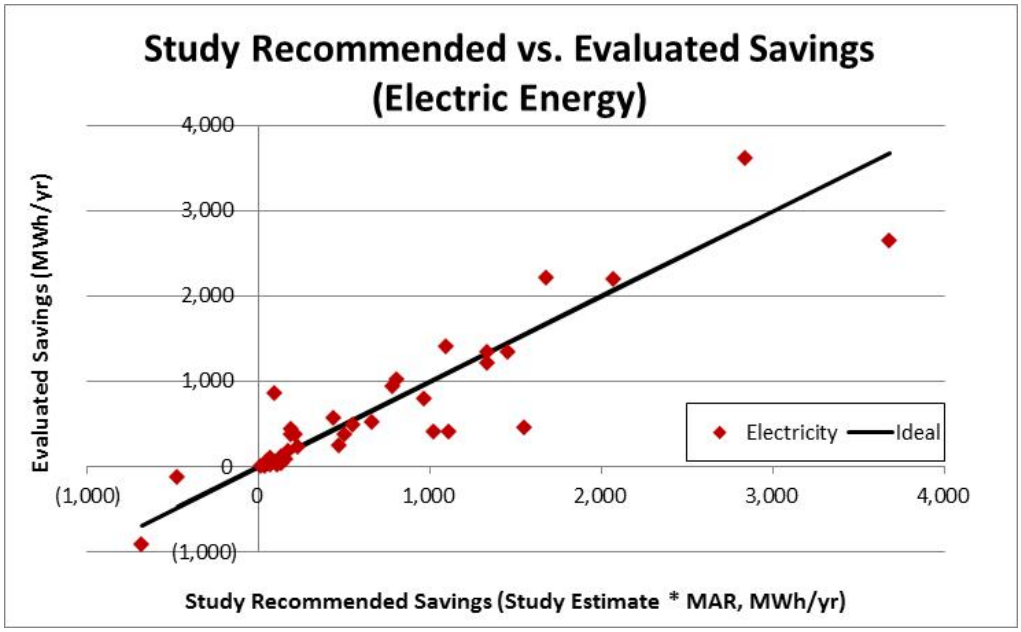
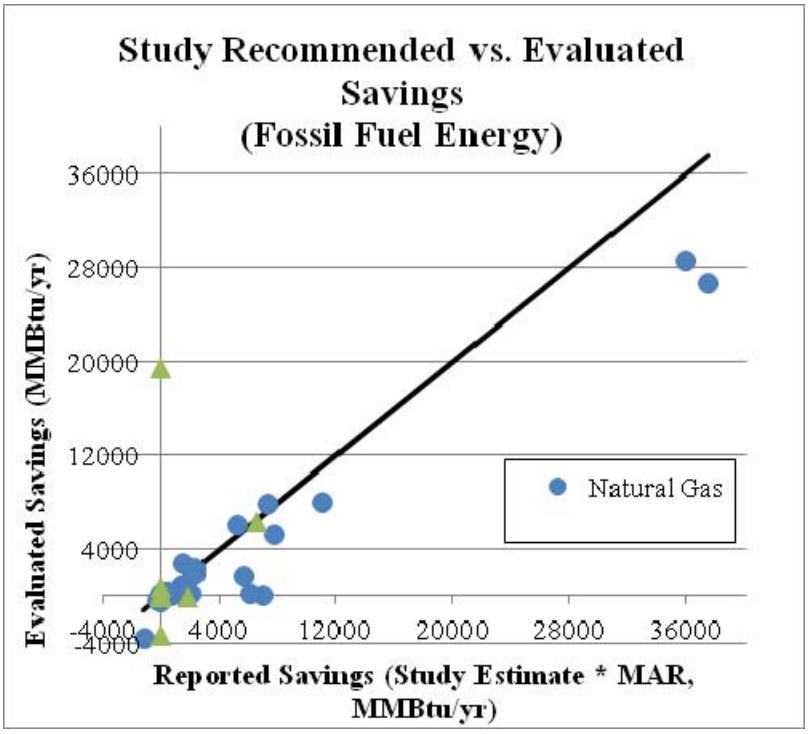


Figure 3-11 shows the same relationship of recommended and evaluated savings for fossil fuel efficiency measures. The three CHP projects are omitted because they each impact use of multiple fossil fuels, some positive and some negative, and this graph does not illustrate that dynamic well. The figure also omits the largest natural gas efficiency measure (91,735 MMBtu/yr recommended savings, 84,502 evaluated) for better resolution of the other points. Overall efficiency measure realization rates varied less for natural gas than they did for electricity.

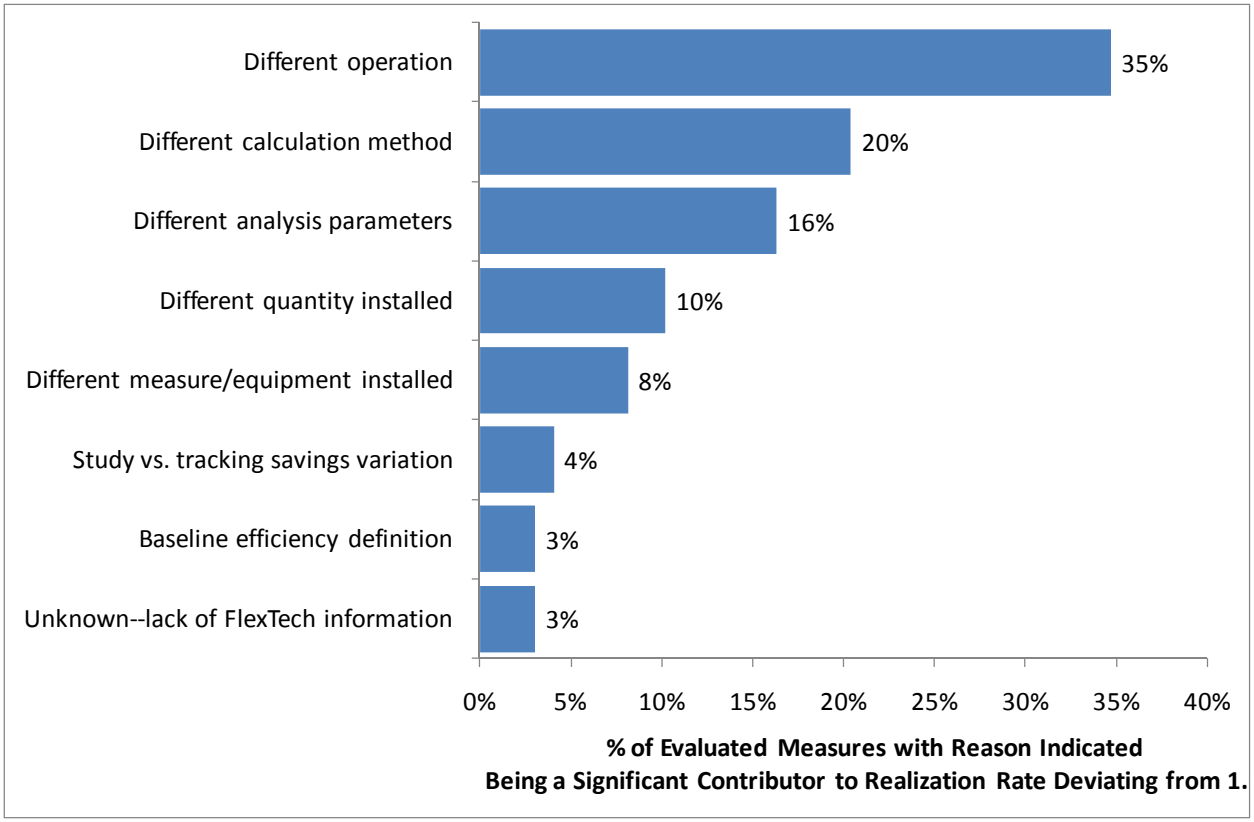
Figure 3-11. Recommended and Evaluated Fossil Fuel Savings, SRR Efficiency Sites Excluding One Largest



3.2.2 Site-Specific Reasons for Energy Savings Realization Rate Deviations from 1.0

The on-site survey evaluated a total of 110 measures. Evaluators calculated savings that deviated from the study-estimated energy savings by more than +/-10% for eighty-six of the measures. For each of those measures the study's lead evaluation engineer categorized up to two major reason(s) for such deviation. The chart in Figure 3-12 illustrates the major reasons for the evaluation results to deviate more than +/-10% when compared with the recommended savings as recorded in NYSERDA's tracking system.

Figure 3-12. Reasons Evaluated Savings Deviated from Study Estimates



Evaluators found that differences in operation (e.g., schedules, hours per year) between predicted and evaluated calculations were the biggest contributing factors. They occurred over half of the time there was significant deviation. No other single factor was responsible for as much as one third of the measures. "Different calculation methods" (e.g., simulation modeling versus metered equipment operation) and "different analysis parameters" were the next largest contributors to the deviations. The latter deviation category indicates that the evaluators made an assumption or used parameters (such as a unit conversion factor) different from the FlexTech analyst. The "study versus tracking" reason indicates that evaluators found differences between the FlexTech study report values and the NYSERDA tracking system savings. It is interesting to note that differences in "baseline efficiency definition," a relatively common reason for deviation in many evaluations, was not a significant contributing factor.

3.2.3 Program-Level Peak Demand Savings Realization Rate

FlexTech study authors estimate demand savings based on many different definitions: Average billed demand savings over 12 months, cumulative billed demand savings per year, super-peak demand savings, summer demand savings, etc. Some studies omit the parameter entirely and estimate economic benefits to the customer based on an average energy rate that included the demand component of the bill.

The previous evaluation perceived the weakness of these estimates and determined FlexTech Program-typical demand savings as a function of energy savings[39]. While the tracking system retains the study estimates, the Program uses the evaluated factor, 5,954 kWh/yr savings per on-peak summer coincident kW of demand savings, to estimate and report program-level impact instead of the study estimates. This evaluation used the factor-based kW savings estimate as "recommended and implemented" savings.

Evaluation analysts calculated site-specific demand savings for the sample using the summer peak coincident demand period definition and used this demand estimate as the evaluated demand estimate, the numerator of the demand SRR ratio. Analysts used the study-recommended annual electricity savings divided by the prior evaluation demand factor to calculate the study-recommended demand savings. Dividing the two resulted in a program-level realization rate of 0.85.

3.2.4 Program-Level Savings Realization Rates

Table 3-4 and Table 3-5 present the electricity-related SRR estimates and associated sampling statistics after weighting the above results. As required by the plan the SRRs are reported separately by region of the state. The overall electric SRRs are 0.92 on energy and 0.86 on summer peak coincident demand. Relative precision is 3.7% on energy and 12.5% on demand. The larger relative precision range on demand was expected, as the program estimates savings based on study energy savings without study-specific demand considerations.

[39] Nexant, M&V Evaluation Technical Assistance Program, prepared for NYSERDA, June 2007.

Table 3-4. SRR Results Summary, Electricity, By Region and Size

Studies	Total Number of Studies	Number of Completed Sites in SRR Sample	Electric Energy		Summer Coincident Peak Demand	
			SRR	SRR Relative Precision	SRR	SRR Relative Precision
Upstate						
Small studies not in sample[1]	35					
Medium and large study sample	70	36	0.87	6.8%	0.87	27.5%
Extra large CHP certainty selection study	1	1	1.12	0.0%	0.62	0.0%
Total	106	37	0.94	6.3%	0.80	29.9%
Downstate						
Small studies not in sample [1]	6					
Medium and large study sample	8	6	0.79	3.1%	1.15	0.5%
Extral large CHP certainty selection study	1	1	1.05	0.0%	0.80	0.0%
Total	15	7	0.87	2.9%	1.05	0.6%
Statewide						
Small studies not in sample	41					
Medium and large study sample	78	42	0.83	4.1%	1.00	10.8%
Extra large CHP certainty selection studies	2	2	1.10	0.0%	0.67	0.0%
Total	121	44	0.92	3.7%	0.86	12.5%

[1] SRR includes 1.0 SRR for small studies not in sample (constituting less than 3% total energy savings).

The Program historically has reported impact based on separate SRRs for energy efficiency measures and for CHP measures. **Error! Reference source not found.** presents the electricity-related results in this fashion.

Table 3-5. SRR Results Summary, Electricity, by Efficiency and CHP Study Types

Studies	Total Number of Studies	Number of Completed Sites in SRR Sample	Electric Energy		Summer Coincident Peak Demand	
			SRR	SRR Relative Precision	SRR	SRR Relative Precision
Efficiency studies	118	41	0.86	5.8%	0.82	23.1%
CHP studies	3	3	1.00	0.0%	0.88	0.0%
Total	121	44	0.92	3.7%	0.86	12.5%

Table 3-6 shows the overall natural gas SRRs and relative precision for efficiency-focused studies.

Table 3-6. SRR Results Summary, Natural Gas Efficiency

Region	Natural Gas Efficiency	
	SRR	SRR Relative Precision
Upstate medium and large sample	0.77	4.0%
Downstate medium and large sample	0.80	0.0%
Total medium and large sample	0.77	4.0%

Some efficiency studies recommended measures that save fossil fuels other than natural gas. Natural gas impact was the focus of the evaluation but when studies included measures with non-gas impacts evaluators calculated them. After comparing natural gas and non-natural gas realization rates and considering the natures of the technologies involved, evaluators find it reasonable to expect that the realization rates for fossil fuels other than natural gas are likely to be similar to that of natural gas.

For CHP studies the program tracking system did not classify any of the fossil fuel impact as being natural gas; thus there is no natural gas SRR. As noted earlier, the tracking system stores a single fossil fuel net impact value and a single fuel type per measure. All three CHP studies in the SRR sample projected impact for multiple fossil fuels; fuels other than natural gas were selected for the tracking system labels. NYSERDA reports aggregate fossil fuel impact to the DPS. Given this tracking system limitation and reporting requirement, aggregate CHP SRR across all fuels is important. Table 3-7 presents the SRR results for all non-electric fuels associated with the three CHP studies.

Table 3-7. SRR Results Summary for Non-Electric Fuels in CHP Studies

Eval Site ID	Recommended (MMBtu/yr)[1]			Evaluated (MMBtu/yr)				Combined Fossil Fuel Savings Realization Rate
	Fossil Fuel Impact	Labeled Fuel Type	Other Fossil Fuels Affected	Natural Gas	Oil	Steam	Total Fossil Fuel	
1	(815,139)	Oil	Natural gas, coal	(664,711)			(664,711)	0.82
7	(159,775)	Steam	Natural gas, oil	(235,317)			(235,317)	1.47
37	(170,931)	Steam	Natural gas	(172,927)		30,839	(142,088)	0.83
Total	(1,145,845)			(1,072,955)	0	30,839	(1,042,116)	0.89a

[1] Recommended values from tracking system. Values in the TA studies anticipate multiple fuel impacts and some differ in magnitude. All measures were adopted at 100% MAR.

[a] The Total SRR calculation incorporates sample weighting factors not shown; therefore it is not the same as the total evaluated / recommended impact for the three studies.

The key SRR analysis findings are as follows:

- The overall evaluated electric energy SRR is 0.92. The electric energy SRR is 1.0 for CHP and 0.86 for efficiency studies.
- For summer coincident peak demand the overall SRR is 0.86 and influence of CHP is similar with an SRR of 0.88 for CHP and 0.82 for efficiency measures.
- The SRR for natural gas efficiency measures is 0.77 and is expected to be similar for other fossil fuels. There is no CHP natural gas SRR. The CHP all-fossil fuels' SRR is 0.89.
- Evaluators found that operating differences, calculation methods, and analysis study parameters were the largest contributors to the deviations observed between the evaluation and FlexTech study results.

The prior evaluation derived a factor of 5,954 study-recommended kWh per kW of demand savings. This evaluation found essentially the same ratio: 5,971 study-recommended kWh per kW of evaluated demand savings, which is within 1% of the prior value. The ratio of evaluated kWh per evaluated kW demand savings was 4,958. These ratios exclude the outlier CHP studies, which each run 8,760 hr/yr and consequently inflate the ratio to over 6,300 kWh per kW when included.

Error ratio is the statistical parameter that measures the site-to-site variance in SRRs. The overall error ratio excluding the outliers is 0.64 on the electric energy SRR and 0.48 on the natural gas SRR. The error ratio is not an informative parameter for demand savings in this evaluation. The SRR sample design assumed an error ratio of 0.60.

3.2.5 Comparison with Prior Evaluation Results

The 2007 evaluation computed a revised SRR of 1.01 on electric and non-electric energy savings.[40] Because this estimate constitutes a revision of the 2005 estimate due to methodological issues rather than a subsequent analysis, the 2005 result is not relevant. It did not calculate demand SRRs. The prior result is higher than this evaluation's overall 0.92 SRR. There are no definitive reasons for the discrepancy between the prior study and current result. This evaluation examined more facilities and did so with greater engineering rigor per study than the prior evaluation could afford to do.

3.3 NET-TO-GROSS EFFECTS

This section covers the components of the impact evaluation that are required to produce the FlexTech Program: free ridership (FR) and spillover (SO) rates and net-to-gross (NTG) ratio. It also provides a discussion of the supplemental information collected through the surveys used to adjust and comparisons that help estimate FR, participant inside SO, and participant outside SO outcomes. The approach and methods to estimate energy savings induced by the FlexTech Program are described in Section 3.6.1, and the survey dispositions are provided in Section 2.6.7.

3.3.1 Free Ridership

As discussed in Section 3.6.1, construct validity is the area of greatest concern regarding estimating free ridership. Like the prior evaluation, FR in this study is addressed through the use of multiple inquiries and consistency checks of a participant's responses. The series of "direct" self-report FR questions was designed to elicit explicit estimates of FR from the respondents. A participant's FR estimate was adjusted when the two direct FR measurements conflict with questions regarding the level of program influence on their adoption decisions. Further adjustment of FR occurs for some sites due to supplemental data collected as part of the enhanced method of this study.

[40] Ibid.

This section is divided into three parts. The first part covers the direct query for both end users and TA service providers. The next section discusses the inclusion of the supplemental data from the on-sites and surveys of additional decision-makers, and the third section integrates the previous two and also incorporates the two sites previously included in the 2008 impact evaluation of the largest savers in the C/I sector.

Direct Query

The initial survey questions relate to whether the end user would have obtained an audit without the Program and the likelihood that it would have been of similar quality as that provided through their participation in the FlexTech Program. The TA service providers were also asked for their opinions about the likelihood that the end user would obtain a similar quality audit outside of the Program. A little over one-third of participating end users (nineteen) reported that it was likely they would have obtained any audit, as presented inTable 3-8.[41] Among this group, most (60%) reported that they would have obtained an audit of similar quality to the one provided through the FlexTech Program. Among end users who reported any chance of obtaining an audit on their own, only 20% would be likely to get an audit of similar quality to that of the Program. From the perspective of the TA service providers, 17% of all end users would have obtained a similar quality audit if they had not been participants of the FlexTech Program, as shown in Table 3-8. Thus, the estimates from the end users and service providers are reasonably close. Figure 3-13 illustrates the same graphically.

Table 3-8. Likelihood of Conducting Energy Audit, and Energy Audit of Similar Quality, without the Program

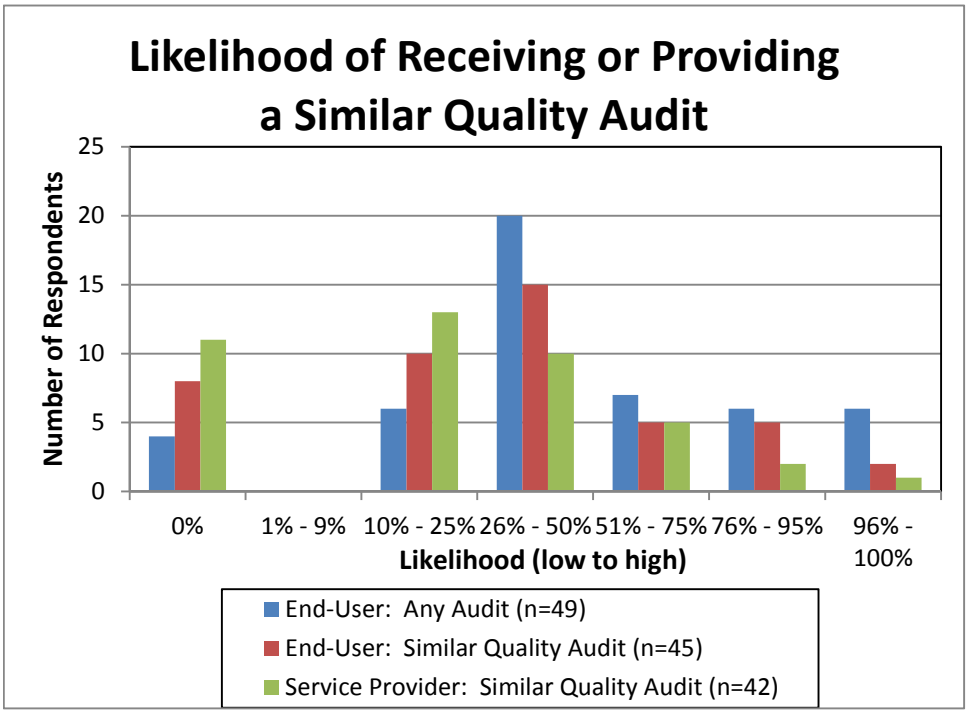
Likelihood Percent	End User		Service Provider
	Any Audit (n=49)	Similar Quality Audit[1] (n=45)	Similar Quality Audit[2] (n=42)
0%	4	8	11
1% - 9%	0	0	0
10% - 25%	6	10	13
26% - 50%	20	15	10
51% - 75%	7	5	5
76% - 95%	6	5	2
96% - 100%	6	2	1

[1] This question was only asked if the response to "Any Audit" was greater than 0%. Four end users were not asked this question due to the survey response skip patterns. These four are included in the totals for calculating the percentages cited in the report.

[2]Four TA service providers stated "Don't Know." These four are included in the totals for calculating the percentages cited in the report.

[41] These end users responded that the likelihood they would have obtained an audit without the Program was greater than 50%.

Figure 3-13. Likelihood of Conducting an Energy Audit of Similar Quality, without the Program



When FlexTech service providers conduct audits outside of the Program, they report that all of these audits are at least as comprehensive as those provided through the Program, as shown in Table 3-9. They estimated that a quarter of the audits are more comprehensive than the audits done through the Program.

Table 3-9. Respondent Direct Free Ridership Estimate (1 of 2)

Non-Program Audit as Compared to FlexTechAudit	Site Information as Provided by Service Providers (n=46)	Percent of Total
More comprehensive	11	24%
Similar in comprehensiveness	35	76%
Less comprehensive	0	0%

The results of the direct FR inquiry are presented in Table 3-10. Respondents were asked to estimate the savings associated with the Program in comparison to what they would have done without the Program in two ways:

- Make a best estimate of the proportion of savings achieved that were directly associated with program participation
- Provide an upper and lower bound on the best estimate

The best estimate of savings associated without the Program was 41% according to the end-use participants and 25% according to the TA service providers.

Table 3-10. Respondent Direct Free Ridership Estimates

	Weighted Average Direct FR Response by End Users (n =47)	Site Weighted Average Direct FR from Participating Service Providers (n=46)
Minimum share of energy savings achieved without Program[1]	31%	18%
Maximum share of energy savings achieved without Program[2]	50%	33%
Best estimate of share of energy savings achieved without Program[3]	41%	25%

[1] Five end users and nine TA service providers responded with "Don't Know" for this item.

[2] Six end users and nine TA service providers responded with "Don't Know" for this item.

[3] Six end users and nine TA service providers responded with "Don't Know" for this item.

Over one-third of the sites had completed surveys from both the site end user and its TA service provider. The prior method derived a weighted average of the end-user-reported FR and averaged that with the weighted average of the TA service provider-reported FR for those same studies. In this evaluation, the end-user and TA service provider FR estimates were averaged at the site level for those sites with estimates from both market actors. For sites with only a single interview with the end user, the site FR is based on the end users' responses. The FR results correspond well, within a narrow range of 28% for sites with only TA service provider responses, 30% for matched sites, and 31% for sites with only end-user responses. Table 3-11 presents these results along with the number of sites in each group.

Table 3-11. Weighted Free Ridership for Sites with Both End-Use and Service Provider Responses v. Free Ridership from Non-Matched Data[42]

Actor Group(s)	Free Ridership Estimate	Number of Sites (n = 55)
End user and matched service provider	30%	20a
End user with no available match	31%	18
Service provider with no available end-user match	28%	17

[a] There were two respondents for each project: the end user and their TA service provider.

Enhanced FR Estimates

To incorporate the additional data from the on-sites and surveys of key decision-makers, sites were classified into three groups according to the data available, and the FR estimate was adjusted as necessary for each site. FR estimates were calculated for each of the three groups and then combined to determine the survey-based FR estimate. The three groups are defined below:

- Group 1: All sites with no supplemental data (fifty-five sites)
- Group 2: Sites with supplemental data from the on-site survey (twelve sites)
- Group 3: Site with data collected from additional decision-makers and from the on-site survey (one site)

[42] This table does not incorporate information gathered from surveys of other decision-makers or site-specific data collected by the engineers.

It was not possible to obtain supplemental data for most sites; therefore most sites fell into Group 1. For these sites, the FR was derived from the direct query results without any further adjustments. Group 2 and Group 3 sites with supplemental on-site data were defined by meeting one of the following two criteria: (1) the on-site data indicated that the firm had a formal written energy efficiency policy and/or (2) the on-site survey score was estimated to reflect the condition and/or efficiency rating for the non-program equipment at a given site.

The free-rider rate, as reported in Table 3-12, is 31% for the first group, 25% for the second group, and 28% for the third group. Group 3 was distinguished by the use of the enhanced methodology in which on-site data was applied to corroborate self-report survey data or to make adjustments for discrepancies between the self-reports and the evidence found at the site.

Table 3-12. Survey-Based Free Ridership Estimates

Analysis Group	Total Site MMBtu	# of sites[1] (n = 67)	Savings Weighted FR
Group 1	754,590	55	31%
Group 2	540	1	25%
Group 3	144,514	10	28%
Final survey-based FR	**899,644**	**--**	**31%**

[1] This value is based on the number of sites in each group in the analysis. The sites are only counted once for combined group analysis.

The FR method used for this FlexTech impact evaluation, as described in Section 3.6, relies largely on the replication of the prior method FR survey questions and algorithm to estimate the participant-specific FR factors with several enhancements. This approach provided sufficient information to perform a strict replication of the prior FR method, allowing us to conduct a valid comparison of the prior method to the enhanced method using the data collected for this evaluation.

The FR rate and NTGR used in NYSERDA's quarterly and annual reports is a blended rate to reflect NYSERDA's approach of reporting cumulative savings. Thus, the blended rate incorporates the findings from previous evaluations. The comparison reported here is strictly between the prior causality evaluation and this impact evaluation to ensure a valid comparison.

Strictly using the prior method with the data collected in this evaluation produces an FR estimate of 36%[43] (see Table 3-13). The FR method from this study's enhanced method produces an FR estimate of 31%, indicating that the FR would have been overstated without incorporating the additional information from the on-sites and the additional decision-makers.

Table 3-13. Comparison of Prior to Enhanced Free Ridership Method

	Free Ridership Estimate
2006 prior method[1] FR estimate (with current data)	36%
2011 impact evaluation enhanced method	31%a

[1] Source: NYSERDA *Technical Assistance Program Market Characterization, Market Assessment and Causality Evaluation, Final Report*, Prepared byQuantec, LLC and Summit Blue Consulting, LLC.June 2007.

[a] This FR estimate includes all the survey-based FR estimates with all enhancements, but does not include the two projects with FR estimates from the 2008 Largest Energy-Saving Projects impact evaluation. Both FR estimates in this comparison use the same data and are fully comparable.

[43] This comparison only included the FlexTech projects with survey results and not the two largest sites. For this comparison, only the sites surveyed for this evaluation were included to ensure that the comparison was valid.

Overall FR Estimate

The FR for two studies had been previously determined as part of the 2008 impact evaluation of its largest energy-saving studies among the commercial and industrial programs. The studies in that evaluation received in-depth gross and net savings methods on a site-by-site basis. The FR from that specialized study was incorporated into the results in this study for these two sites. The survey-based FR estimate as described above was applied to the population of program participants with adopted measures and then the FR estimates for those two sites from the prior impact evaluation were incorporated into the final program FR estimate. This process resulted in a final program FR estimate of 32% as shown in Table 3-14.

Table 3-14. Free Ridership Rates

	Free Ridership Rate
Survey–based	31%
Largest energy-savings studies (2) (evaluated in 2008)	75%
Overall free ridership rate	**32%**

For the most part, the FR rates estimated through the various survey-based methods are reasonably close, ranging from 25% from the TA service provider direct responses to 31% for the enhanced FR estimate. The one exception is the end-user direct response estimate of 41%. The final FR estimate includes the various types of program influence, covering the likelihood of taking action to install high efficiency systems, the proportion of installs that were high efficiency versus standard efficiency, and when the participant would have taken action versus when it occurred through the Program. The directional difference between the simplistic direct estimate from the end users and the final estimate could be due to the multiple ways in which a program can influence the energy efficiency decision and that not all of these are considered when an end user is asked this general question.

3.3.2 Participant Spillover

Spillover (SO) is defined as energy efficiency savings that are induced by the Program, but are not directly included in the gross program savings. The possible spillover components include:

- Participating end-user inside (within site) spillover
- Participating end-user outside (within firm, other location) spillover
- Non-participating end-user outside spillover

The sources of the spillover estimates for this study have been provided by the following sources:

- Participating end-user inside SO – based on end-user surveys
- Participating end-user outside SO – based on participating TA study provider surveys
- Both participating and non-participating end-user SO – two sources: participating and non-participating TA Study Providers

The non-participant spillover (NPSO) is the sum of two estimates of spillover, one estimated in this study from the participating study provider surveys and the second estimated in another study which is discussed in the next section.

Participant Inside Spillover

Inside SO is the program-induced energy efficiency savings that were not included in the Program at the same site/building as the program savings. The inside SO estimate from this evaluation is based on responses from the end-user survey. The TA Service Providers may be able to answer questions concerning whether the participating firm did additional energy efficiency actions at the same building as was included in the TA study

that the TA Service Provider conducted. But the accuracy for this estimate would be expected to be much greater coming from the firm itself.

Over one-quarter of the participating end users state their firm installed or adopted additional energy efficiency actions at the site that participated in the FlexTech Program and the Program was at least partially the cause of these savings generated outside the Program. Most of these (85%) had electric spillover and about half had natural gas spillover. These numbers and the number of survey respondents providing these responses are shown in Table 3-15.

Table 3-15. End Users Reporting Inside Spillover

Participant Inside Spillover	Yes (Percent, n)	No (Percent, n)	Don't Know (Percent, n)
Inside SO (n = 47)	28% 13	66% 31	6% 3
Electric inside SO (n = 13)a	23% 11	15% 2	0%
Average extra electric savings	2,670.98 MMBtus		
Natural gas inside SO (n = 13)a	15% 7	46% 6	0%
Average extra natural gas savings	6,275.84 MMBtus		

[a] Only participants that said yes they had spillover in the building/study that participated in the Program were asked the follow-up questions concerning electricity and natural gas spillover.

Participants reporting SO were asked about the magnitude of the energy savings relative to the program induced savings at the site, by fuel type. From these answers, the SO savings were estimated for electricity and for natural gas for each participant with inside SO. The sum of the electric SO savings divided by the total electric program savings for all survey respondents provides the electric inside SO rate[44]. The natural gas inside SO rate is calculated in the same way. The sum of SO MMBtus divided by the sum of all program MMBtus from all survey respondents (with electric savings converted using source MMBtus) is the overall inside SO rate. This evaluation found participant inside SO for the FlexTech Program to be 4%.

Participant and Non-Participant Outside Spillover

Participant outside SO is program-induced savings from energy efficiency installations and actions not included in the Program and occurring at the firm's other buildings for participating building owners. Non-participant outside SO is program-induced savings from energy efficiency installations and actions occurring at non-participant sites.

The outside SO was estimated from the survey of participating TA Service Providers and covers a broad definition of outside SO and includes SO at participant and non-participant end-user sites. The prior FlexTech causality study also based the outside SO estimate on a survey of service providers. However, this SO estimate only partially captures non-participant SO, requiring additional input from non-participating TA Service Providers to fully capture SO, as discussed in the next section.

[44] Electric savings is converted to source MMBtu multiplying the kWh savings by 0.0099492.

Over two-thirds of TA Service Providers report outside SO, installing or assisting in the adoption of additional energy efficiency measures in non-program buildings and sites of both participating and non-participating end users. Almost all of these had both electric and natural gas spillover, as can be seen in Table 3-16.

Table 3-16. TA Service Providers Reporting Outside Spillover

Outside Spillover	Yes (Percent, n)	No (Percent, n)	Don't Know (Percent, n)
Outside SO (n = 27)	67% 18	33% 9	0%
Of the 67% (18) of participating vendors that had outside spillover:			
Electric outside SO (n = 18)a	94% 17	0% 0	6% 1
Average electric savings	45,159.06 MMBtus		
Natural gas outside SO (n = 18)a	89% 16	11% 2	0%
Average natural gas savings	8,296.33 MMBtus		

[a] Only TA Service Providers that said yes their participation in the Program led them to install additional measures in non-program buildings and studies, i.e., reported outside SO, were asked the follow-up questions concerning electricity and natural gas spillover.

The TA Service Providers were then asked about the number of buildings with outside SO and the percentage of electricity and natural gas savings from these additional measures as compared to the average electric and natural gas savings they see within their Program participants. The responses from these steps allow for the calculation of outside SO savings estimates for electricity and natural gas. As described above for participating end users, SO savings are combined across fuel type by converting electricity SO savings into MMBtus.

Current service provider respondents reported outside SO that results in a 30% outside SO rate for both participating and non-participating end users.

Other Estimates of Spillover for Comparison

For comparison purposes, end users were also asked about outside SO in facilities other than the selected site. In order to avoid double counting with the TA service provider based estimates, these results were not explicitly included in the SO calculations.

Over one-third (34%) of participating end users report having outside SO. While this result is higher than the technical service provider reports of 22% of end users with outside SO (Table 3-17), it could indicate that end users may be using other firms to implement additional non-program installations.

Table 3-17. Participating TA Service Provider Claims End User Had Outside Spillover

End User Had Spillover	Percent	Number of Respondents (n=27)
Yes	22%	6
No	56%	15
Don't know	22%	6

Spillover Conclusions

Table 3-18 provides a summary of the SO estimates calculated from the NTG surveys in this evaluation. Participating end users report inside SO as a 4% rate. The outside SO rate for participating end users is 2%. The TA Service Providers report an outside SO rate of 30%.

Table 3-18. Inside and Outside Participant Spillover Estimates

Type of Participant Spillover	Number of Respondents (Used to Derive Spillover Rate)	Spillover Rate
End user inside SO	14	4%
Participating service provider outside SOa	17b	30%c

[a] The final net-to-gross ratio uses the TA service provider outside SO as it is a broader definition of outside SO. The end-user outside SO is not used in the NTG as including it and service provider outside SO could result in double counting.

[b] There were only seven respondents with relatively complete data.

[c] This estimate is highly uncertain due to the low number of complete responses and high variability among the responses.

The estimate of outside SO from the TA service providers is highly uncertain due to the low number of complete responses and the high degree of variability in the responses. It is clear that new methods are needed for future evaluations. Since significant outside SO is reported by the TA service providers, basing an estimate on these responses is preferable to setting the outside SO to zero or selecting an arbitrary value.

3.3.3 Indicators of Non-Participant Spillover Provided by TA Service Providers

This evaluation did not include estimating the NPSO rate by non-participating TA Service Providers. NYSERDA's effects on non-participants in the commercial and industrial markets may be due to the FlexTech Program, NYSERDA's large direct assistance program (Existing Facilities), its sector efforts, and/or marketing and technology development support through research and development. Consequently, NYSERDA fielded a NPSO study across the C/I sectors rather than program by program. The NPSO estimate[45] from the 2007 study was used to determine the overall net-to-gross ratio for the FlexTech Program.

This evaluation did, however, include survey questions for the TA Service Providers to gain their perspective on the use of energy efficient practices among non-participating firms and test whether their views provided corroborating evidence for using SO from the previous C/I NPSO study. These results support the finding of NPSO from the previous evaluation.

More participating TA service providers say they have seen an increased use of energy efficiency measures among non-program engineering firms then those who reported no change (and only one respondent said they saw decreased use of energy efficiency measures among non-program firms). Table 3-19 provides these response rates, showing that 41% of the respondents reported an increase while 37% observed no change in the use of energy efficiency measures.

[45] "The MCAC Team used detailed telephone surveys with non-participating C&I end-use customers to reassess the previously estimated non-participant spillover value and generate information regarding additional items of interest to NYSERDA including familiarity with energy efficiency measures and equipment and barriers that prevent more high efficiency measures and equipment from being installed in the market. The surveys were designed around assessment indicators identified in the C&I sector-level logic model report and incorporated questions used in prior research efforts to examine longitudinal findings in terms of these indicators." Summit Blue, *Commercial and Industrial Market Effects Evaluation*, p. 2-2.

Table 3-19. Perceived Changes in Use of EE Measures among Non-Program Firms

	Sites with Service Provider Respondents (n=46)	Percentage
Increased use	19	41%
Decreased use	1	2%
No change	17	37%
Don't know	9	20%

Of those reporting an increased incorporation of energy efficiency, one-third said energy efficiency measures being incorporated by non-program firms increased a lot and two-thirds said it increased a little, as shown in Table 3-20.

Table 3-20. Magnitude of Change among Non-Program Firms for Firms that Perceived an Increase in Energy Efficiency

Degree of Change	Sites with Service Provider Respondents (n=20)a	Percentage
Increased a lot	6	30%
Increased a little	13	65%
Don't Know	1	5%

[a] Twenty-six respondents were not asked this question due to skip pattern instructions.

Among those who were asked, most of the respondents credited the Program for influencing the increases observed, with over two-thirds indicating that the FlexTech Program was a major influence (see Table 3-21).

Table 3-21. FlexTech Influence on Market Change Regarding Use of EE Measures

Degree of Influence	Sites with Service Provider Respondents (n = 19)a	Percentage
Major influence	12	63%
Minor influence	4	21%
Not a factor	1	5%
Don't Know	2	11%

[a] Twenty-seven respondents were not asked this question due to skip pattern instructions.

3.3.4 Final Net-to-Gross Ratio

The FR rate and SO rate are combined to produce a net-to-gross (NTG) ratio that is applied to evaluation-estimated gross savings to produce net savings.

$$\textit{Net - to - Gross Ratio (NTGR)} = 1 - \textit{Free ridership factor} +$$

$$\textit{Inside participant spillover factor (from the participating end - user)} +$$

$$\textit{Outside participant spillover factor (from the participating TA service providers)} +$$

$$\textit{Non-participant spillover factor (from the C/I non-participant spillover study)}$$

The NTGR from this evaluation, and including the latest non-participant spillover factor, was calculated as follows:

$$NTGR = 1 - 0.32 + 0.04 + 0.30 + 0.15$$

for a total NTGR of 1.17. (See Figure 3-14)

This is the NTGR that is applied to the evaluated gross savings for the FlexTech Program for this impact evaluation. The NTGR is comparable to the prior FlexTech causality study and changes are in the direction and magnitude of what would be expected[46].

Figure 3-14. Final FlexTech 2007 – 2009 Net-to-Gross Ratio

1	**-**	**32%**	**+**	**4%**	**+**	**30%**	**+**	**15%a**	**=**	**1.17**
Full Gross Savings		Free Ridership		End-User Inside Participant Spillover		TA Service Provider Outside Participant Spillover		Overall C/I Non-Participant Spillover		NTGR FlexTech 2007-2009

[a] Source: NYSERDA *Commercial and Industrial Market Effects Evaluation*, Final Report, Submitted by Summit Blue Consulting LLC and Quantec, LLC.October 2007, page 4.

3.4 NET IMPACT

The net impact calculations combine the results of the three prior sections on MAR, SRR, and NTG, into overall program impact.

The focus of this impact evaluation was to report the program impacts for studies completed from January 1, 2007 through December 30, 2009. Some of the impact will occur after this evaluation is completed.

The natural gas net impact estimates use the natural gas efficiency-based SRR of 0.77 for all studies. No natural gas CHP-based SRR could be calculated and there was only one CHP study completed between 2007 and 2009, for which no natural gas impact was reported.

Table 3-22, Table 3-23, and Table 3-24 estimate both the past impact and projected future impact based on the long-term MAR data and the SRR and NTG factors. Figure 3-15, Figure 3-16, and Figure 3-17 show that measures will continue to be adopted in the future based on the MAR curves presented earlier. The implementation profile has a sliding profile – with the 2007 year studies measures achieving terminal status first followed by the 2008 and then the 2009 studies measures.

[46] Technical Assistance Program Market Characterization, Market Assessment and Causality Evaluation, Final Report, Prepared by Quantec LLC and Summit Blue Consulting, LLC, 2007.

Table 3-22. Past and Expected Long-Term Net Impact, Electric Energy, Studies Completed 2007 – 2009

Parameter	Study Completion Year			Total
	2007	2008	2009	
Study recommended and tracked savings (MWh/yr)	59,239	46,421	40,990	146,651
Electric Energy, MAR prior to 2007	0.01			0.00
Electric Energy, MAR 2007-2008	0.25	0.01		0.10
Electric Energy, MAR 2008-2009	0.09	0.25	0.01	0.12
Electric Energy, MAR 2010-2011	0.13	0.09	0.25	0.15
Electric Energy, MAR 2011-2012	0.15	0.13	0.09	0.13
Electric Energy, MAR 2012-2013	0.01	0.15	0.13	0.09
Electric Energy, MAR 2013-2014	0.04	0.01	0.15	0.06
Electric Energy, MAR 2014-2015	0.00	0.04	0.01	0.02
Electric Energy, MAR 2015-2016		0.00	0.04	0.01
Electric Energy, MAR 2016-2017			0.00	0.00
Electric Energy, MAR, long-term total	0.68	0.68	0.68	0.68
Electric Energy, SRR	0.92	0.92	0.92	0.92
NTG	1.17	1.17	1.17	1.17
Net impact, long-term (MWh/yr)	43,360	33,978	30,003	107,341

Figure 3-15. Past and Expected Long-Term Net Impact, Electric Energy, Studies Completed 2007 – 2009

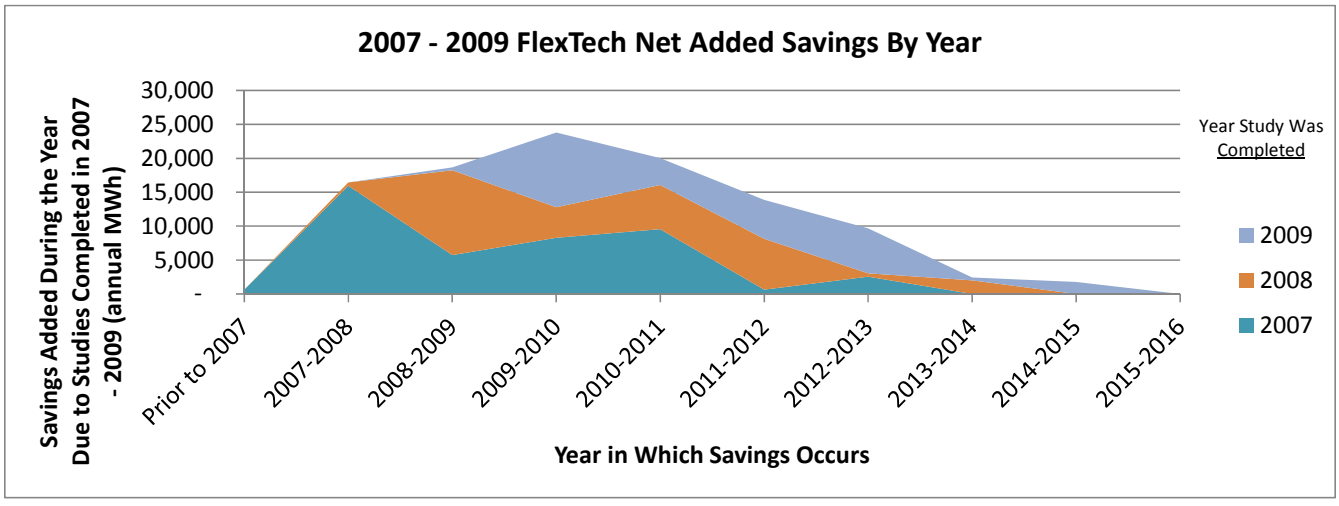
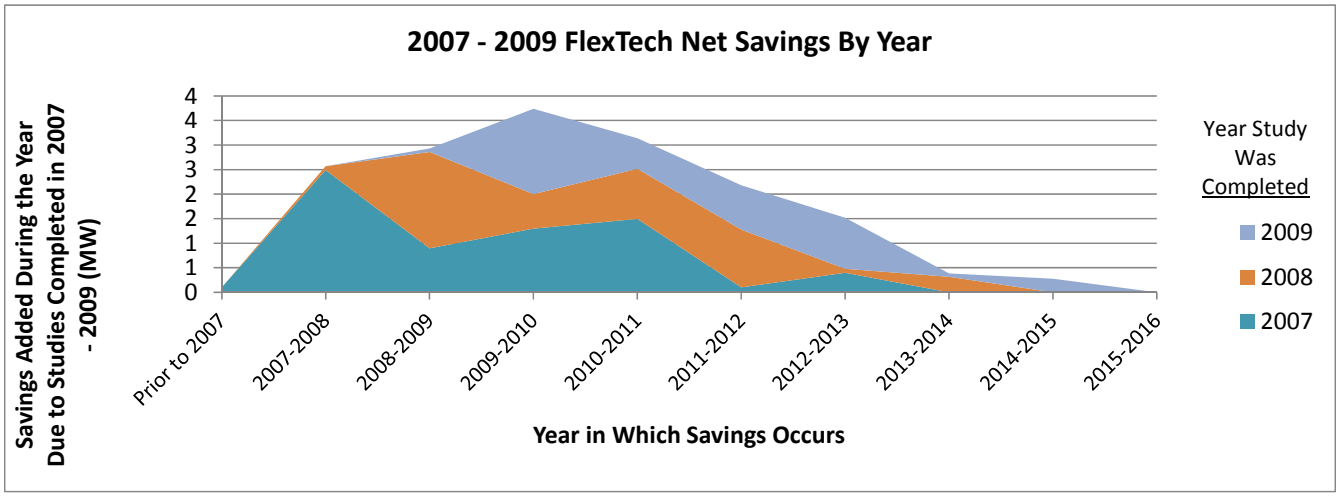


Table 3-23. Past and Expected Long-Term Net Impact, Electric Demand, Studies Completed 2007 – 2009

Parameter	Study Completion Year			Total
	2007	2008	2009	
Reported savings (MW[1])	9.9	7.8	6.9	24.6
Electric Demand, MAR prior to 2007	0.01			0.00
Electric Demand, MAR 2007-2008	0.25	0.01		0.10
Electric Demand, MAR 2008-2009	0.09	0.25	0.01	0.12
Electric Demand, MAR 2010-2011	0.13	0.09	0.25	0.15
Electric Demand, MAR 2011-2012	0.15	0.13	0.09	0.13
Electric Demand, MAR 2012-2013	0.01	0.15	0.13	0.09
Electric Demand, MAR 2013-2014	0.04	0.01	0.15	0.06
Electric Demand, MAR 2014-2015	0.00	0.04	0.01	0.02
Electric Demand, MAR 2015-2016		0.00	0.04	0.01
Electric Demand, MAR 2016-2017			0.00	0.00
Electric Demand, MAR, long-term total	0.68	0.68	0.68	0.68
Electric Demand, SRR	0.86	0.86	0.86	0.86
NTG	1.17	1.17	1.17	1.17
Net impact, long-term (MW)	6.8	5.3	4.7	16.8

[1] The reported demand savings were estimated using prior evaluation derived factor of 5,954 study-recommended kWh per kW of demand savings.

Figure 3-16. Past and Expected Long-Term Net Impact, Electric Demand, Studies Completed 2007 – 2009

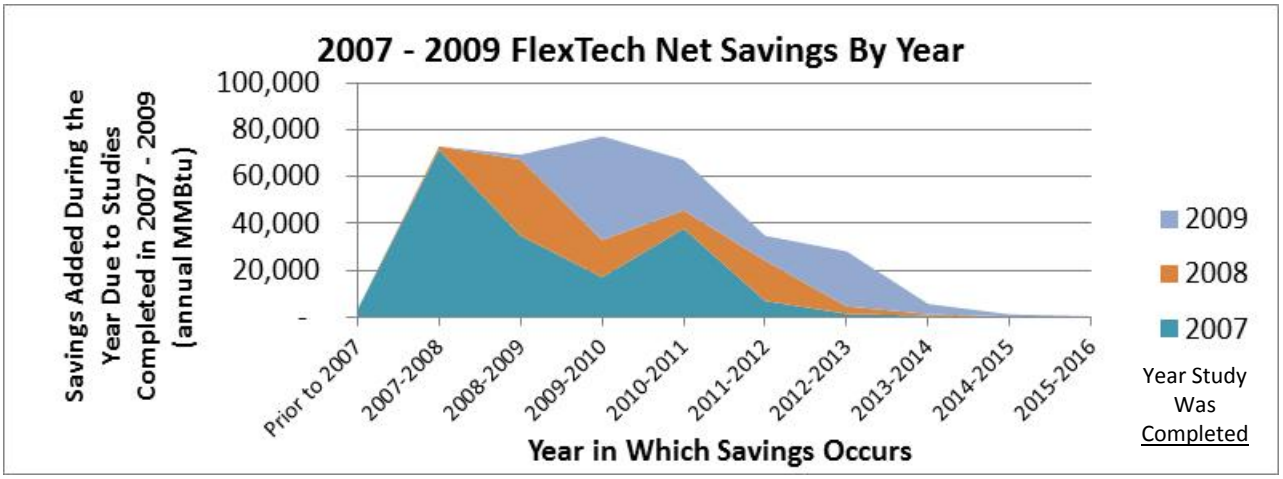


The natural gas net impact estimates use the natural gas efficiency-based SRR of 0.77 for all studies. No natural gas CHP-based SRR could be calculated and there was only one CHP study completed between 2007 and 2009, for which no natural gas impact was reported.

Table 3-24. Past and Expected Long-Term Net Impact, Natural Gas, Studies Completed 2007 – 2009

Parameter	Study Completion Year			Total
	2007	2008	2009	
Study recommended and tracked savings (MMBtu/yr)	447,247	203,570	277,206	928,023
Natural Gas, MAR prior to 2007	0.01			0.00
Natural Gas, MAR 2007-2008	0.18	0.01		0.09
Natural Gas, MAR 2008-2009	0.09	0.18	0.01	0.08
Natural Gas, MAR 2010-2011	0.04	0.09	0.18	0.09
Natural Gas, MAR 2011-2012	0.09	0.04	0.09	0.08
Natural Gas, MAR 2012-2013	0.02	0.09	0.04	0.04
Natural Gas, MAR 2013-2014	0.00	0.02	0.09	0.03
Natural Gas, MAR 2014-2015	0.00	0.00	0.02	0.01
Natural Gas, MAR 2015-2016		0.00	0.00	0.00
Natural Gas, MAR 2016-2017			0.00	0.00
Natural Gas, MAR, long-term total	0.43	0.43	0.43	0.43
SRR	0.77	0.77	0.77	0.77
NTG	1.17	1.17	1.17	1.17
Net impact, long-term (MMBtu/yr)	173,258	78,860	107,386	359,504

Figure 3-17. Past and Expected Long-Term Net Impact, Natural Gas, Studies Completed 2007 – 2009



The net impacts above estimate both historic and expected future adoption associated with the 2007 through 2009 FlexTech studies and are based on the MAR survey of 2003 through 2009 participants.

Table 3-25below summarizes the estimated actual impact of the 2007 through 2009 studies at the time of the MAR survey in the summer of 2010. The net impacts as of 2010 differ from those in the prior tables in that they are historic only, based on a snapshot in time. They correlate with the results for which sampling statistics are reported.

Table 3-25. Net Impact as of Summer 2010, Studies Completed 2007 – 2009

Parameter	Sample Frame Period	Efficiency			On-Site Generation (MWh/yr)
		Electric Energy (MWh/yr)	Electric Demand (MW)	Natural Gas (MMBtu/yr)	
Study-recommended and tracked savings	1/1/07 – 12/31/09	146,651	24.6	928,023	82,755
MAR as of summer 2010	1/1/07 – 9/30/09	0.35	0.35	0.31	0.18
SRR	1/1/06 – 9/30/09	0.92	0.86	0.77	na
Adjusted gross impact		47,222	7.4	221,519	0
NTG	1/1/06 – 9/30/09	1.17	1.17	1.17	1.17
Net impact to date		55,250	8.7	259,177	0

Table 3-26 below summarizes the estimated actual impact of the 2007 through 2009 studies over the long term. The SRR and NTG estimates do not change as they are not considered to be time dependent.

Table 3-26. Net Impact Long Term, Studies Completed 2007 – 2009

Parameter	Sample Frame Period	Electric Energy (MWh/yr)	Electric Demand (MW)	Natural Gas (MMBtu/yr)
Study recommended and tracked savings	1/1/07 – 12/31/09	146,651	25	928,023
MAR long-term	1/1/03 – 9/30/09	0.68	0.68	0.43
SRR	1/1/06 – 9/30/09	0.92	0.86	0.77
Adjusted gross impact		91,745	14.4	307,268
NTG	1/1/06 – 9/30/09	1.17	1.17	1.17
Net impact to date		107,342	16.8	359,504

3.5 NET IMPACT TOWARDS 15 X 15 GOALS

All net impact above is associated with the impact of studies completed in 2007 through 2009. When measuring progress toward the 15% savings by 2015 (15 x 15) goals, NYSERDA accounts for all savings installed January 1, 2007 or later. Table 3-27 summarizes the 15 x 15 net impacts by year.

Table 3-27 uses the current evaluation MAR factors for studies that fall in the range of this evaluation (2007-2009). The current evaluation SRR and NTG factors are used for studies completed in 2007 and later. The prior evaluation SRR (1.01) and NTG (1.17) estimates are used for studies completed prior to 2007. The current evaluation team could not use the prior evaluation MAR estimates for the studies conducted before 2007 due to the fundamental difference (described in 3.4.1) between the prior and current evaluations.

The totals are 248 annual GWh, 39 MW, and over 489,535 annual MMBtu.

Table 3-27. Net Impact Long Term Towards 15 x 15 Goals, Studies Completed 2003 – 2009

Year Savings	Electric Energy (MWh/yr)	Electric Demand (MW)	Natural Gas (MMBtu/yr)
2007	79,401	12.5	136,872
2008	63,519	10.0	119,229
2009	48,066	7.6	94,632
2010	25,707	4.0	68,093
2011	17,445	2.7	34,554
2012	9,669	1.5	27,756
2013	2,440	0.4	6,534
2014	1,765	0.3	1,366
2015	-	-	499
Total	**248,012**	**39.0**	**489,535**

3.6 EARLY REPLACEMENT

The Impact Evaluation Team collected data on the age of replaced equipment and the reasons for replacement. These questions were asked for all of the replacement measures in the study sample, up to five measures per site. The measures were binned into twelve different categories. The complete survey data and expanded definitions of the categories are presented in Appendix I.

3.6.1 Early Replacement Survey Results

The following section provides details on the equipment replacement responses received for three questions.

Question 1: To the best of your recollection how old was the equipment you replaced?

Table 3-28 presents the response frequencies of the answers provided to this question for all measures in the sample. Several trends can be observed from the answers to this question:

- Overall, over 40% of the respondents said that the replaced equipment was over 20 years old. Considering that most of the energy efficiency measures included in this analysis have lives between 15 and 20 years, the data suggests assessing measure life is a complicated issue and that some of these replacements likely should be treated as end-of-life time measures.
- Most of the measures involving relatively extensive capital investment (building envelope, new compressors, new HVAC systems, new motor systems) show a higher percentile (over 50%) of older equipment (over 20 years old) being replaced. Many of these measures replaced equipment that was 30 to 40 years old. This was somewhat expected considering the cost of these measures. However, this is an indication that measure life is a complicated issue and the current measure life used in the programs might be too short.
- The lighting measures tend to have components installed at shorter time intervals, as over 45% of the measures replaced equipment that was under 15 years old.
- The controls measures tend to be more evenly distributed across the age categories, but only 50% of the respondents provided information regarding the age of the equipment.
- Approximately 9% of the measures did not replace older equipment and were installed for other reasons.

Table 3-28. Question 1 (REP1) Response

Measure	Under 10 Years	10 to 15 Years	15 to 20 Years	20 to 25 Years	Over 25 Years	Did Not Replace Older Equipment	Not Yet Installed	Don't Know	Not Applicable	Totals
Building envelope	0	0	0	2	3	2	0	0	0	7
Compressed air leaks	2	0	1	1	2	0	1	1	0	8
Compressed air: other	0	0	4	0	0	2	0	1	0	7
Compressed air: VFD compressors	0	1	1	2	0	0	0	0	0	4
Controls	4	1	1	1	2	4	0	4	0	17
Heat recovery	2	0	0	0	0	1	1	0	0	4
HVAC equipment replacement	2	1	1	1	14	1	0	2	0	22
Insulation	1	1	0	0	0	0	0	2	0	4
Lighting and lighting controls	12	3	2	4	9	1	0	2	0	33
Motors replacement	1	1	1	1	4	0	0	0	0	8
Other	2	3	2	1	8	2	0	5	0	23
VFDs	1	0	0	0	4	1	0	1	0	7
WWTP measures	1	2	0	0	2	0	1	1	0	7
Totals	**28**	**13**	**13**	**13**	**48**	**14**	**3**	**19**	**0**	**151**

Question 2: Which of the following best describes your decision to replace the equipment?

Table 3-29 presents the response frequencies of the answers provided to this question for all measures in the sample. Several trends can be observed from the answers to this question:

- Overall, over 45% of the respondents said that the replaced equipment was working but it was not as efficient as newer models. Several of the categories (compressed air, insulation, HVAC replacement, lighting, WWTP measures) show even higher percentiles for this response, with some as high as 75%. This is somewhat expected considering the fact that many facilities decide to replace operating equipment with more efficient equipment rather than wait until the end of its "nominal" life.
- While the age question revealed that 40% of the replaced equipment was over 20 years old, less than 10% responded that the equipment did not work or needed frequent maintenance. This adds further weight to the indication that measure life is a complicated issue and the current measure life used in the programs might be too short.
- Twenty percent of the respondents said that the equipment worked, but was old and would need to be replaced in the next couple of years. These types of measures may have to be analyzed using the dual baseline methodology in the future[47].

[47] On July 18, 2011, as a part of case 07-M-0548, the NY DPS issued an order approving modifications to the Technical Manual (see Appendix M) that requires the use of dual baseline methodology in special-circumstance replacement projects.

- Approximately 4% of the responded indicated that the equipment worked, but they needed either a smaller or larger system. This response suggests that the measures under this classification may require to be treated as end-of-life but is not necessarily so.

Table 3-29. Question 2 (REP2) Response

Measure	It Was Working, but Not as Effectively as Newer Models	It Was Working, but We Needed a Smaller/ Larger System	Working, but Old and Would Need to Be Replaced in the Next Couple of Years Anyway	It Required Frequent Maintenanc	It Was Not Working	Other	Not yet Installed	Refused	Don't Know	No Answer	Not Applicable	Totals
Building envelope	2	0	0	0	3	0	0	0	0	0	2	7
Compressed air leaks	3	0	1	2	1	0	0	0	0	1	0	8
Compressed air: other	4	0	0	0	0	0	0	0	1	0	2	7
Compressed air: VFD compressors	3	0	1	0	0	0	0	0	0	0	0	4
Controls	5	0	2	1	1	0	0	0	4	0	4	17
Heat recovery	2	0	0	0	0	0	0	0	0	1	1	4
HVAC equipment replacement	11	1	6	2	1	0	0	0	0	0	1	22
Insulation	3	0	0	0	0	0	0	0	1	0	0	4
Lighting & lighting controls	24	1	5	0	0	0	0	0	2	0	1	33
Motors replacement	4	0	2	2	0	0	0	0	0	0	0	8
Other	3	2	10	1	1	1	0	0	3	0	2	23
VFDs	2	1	3	0	0	0	0	0	0	0	1	7
WWTP measures	4	1	0	0	0	1	0	0	1	0	0	7
Totals	**70**	**6**	**30**	**8**	**7**	**2**	**0**	**0**	**12**	**2**	**14**	**151**

3.6.2 Early Replacement Conclusions

The early replacement versus retrofit survey revealed significant trends regarding the decision-making associated with implementing energy efficiency measures based on the age of the equipment:

- A significant portion of the equipment (40%) replaced by the investigated measures was over 20 years old. This might be an indication that some of these measures were end-of-life measures, but also indicates that measure life is a complicated issue and that the measure life used in the energy efficiency programs might be too short.
- While the replaced equipment was old, a significant portion of the respondents (45%) claim that they decided to replace it because although it was working, it was not as effective as newer models would be. This is an indication of the market place practices to maintain the existing equipment rather than replacing with higher

efficiency at the end of its "nominal" life. Also, this clearly indicates that the age of the equipment was not the most important reason for replacing the equipment.

- Technology-wise, the higher capital cost measures, such as building envelope, new compressors, new HVAC systems, and new motor systems show older age of the replaced equipment, but also a very high rate of replacement for energy efficiency purposes.
- A large portion of the lighting and controls measures replaced equipment that was less than 10 years old. This is expected for these types of measures, which tend to have a higher turnaround based on the technology advancements and capital cost. Lighting and controls adjustments are shorter life measures and, as such, are implemented more frequently.
- Considering the equipment reported as working but in need of replacement in the next couple of years, equipment requiring frequent maintenance, and equipment not working, we estimate that 15%-20% of the total number of measures might be regarded as end-of-life measures.

Overall, the survey data shows that there are a significant number of measures that replace old equipment. However, the demarcation line between end of life and early replacement is hard to determine as the participants indicated that the equipment is operating well. Applying a dual baseline for energy savings might also be hard to implement as much of the existing equipment is 15 to 20 years old and could be used for another 15 to 20 years with proper maintenance. The survey clearly indicates that the age of the replaced equipment is predominantly over 20 years.

More research is needed to determine if a dual baseline is to be used when reporting savings for energy efficiency measures. The research should also include collecting information regarding maintenance cycles, determining how the decision to replace equipment rather than repair it is made (financial, efficiency, maintenance, other), and better differentiating between technologies.

3.7 PROGRAM OVERLAP DATA COLLECTION DURING MAR SURVEY

NYSERDA previously contracted a portfolio-level overlap study that computed the percentage of savings associated with the FlexTech Program that also received NYSERDA program installation incentive funding.[48] The primary research method was program database comparison. The exercise was challenging because NYSERDA does not associate a unique customer or premise ID . Thus matches had to be made on names, addresses, project descriptions, and similar parameters. The overlap study positively identified and quantified measure overlap. Thanks to the level of detail available in the portal and the method, there was little chance of error in identifying a project as being an overlap project that actually was not one. The converse is not necessarily true. It is unknown if there were projects not identified as overlapping that did in fact overlap. It was cost-prohibitive to check all non-matching studies for overlap. Thus it is possible that overlap was under-identified and that there is bias in the overlap estimate.

NYSERDA recognized this possibility and requested that telephone interviewers collect overlap data from participants in conjunction with the MAR research. Questions were posed regarding external funding from NYSERDA, from investor-owned utility programs, and from government sources. The response quality was low; about half of the respondents did not know or did not answer whether or not they received installation incentives associated with the FlexTech measures. Even so, the responses provide enough information to identify some possible overlap projects without the database matching approach and enable the next overlap study researchers

[48] *Supporting Documentation for the Estimate of Overlapping Savings for the Technical Assistance Program Memo*, prepared for the New York State Energy Research and Development Authority by Kathryn Parlin, West Hill Energy & Computing and Lori Megdal, Ph.D., Megdal & Associates (M&A), September 2008. According to this study, 19.3% of the FlexTech electric energy savings overlap with the Existing Facilities program and 26.8% with the DG/CHP program. The overlap study also found that 20% of the FlexTech demand savings overlapped with the DG/CHP program.

to investigate whether there is a meaningful number of overlapping projects that database comparison does not reveal.

The information has been delivered to NYSERDA. There are no associated results to report.

Section 4:

CONCLUSIONS AND RECOMMENDATIONS

4.1 CONCLUSIONS

Engineers determined the measure adoption status of 2,452 measures recommended for implementation in FlexTech studies between 2003 and 2009. Following a telephone MAR survey, engineers conducted site visits to determine realization rates of installed projects. Separately, telephone interviews were conducted to determine the program's free ridership (FR) and spillover (SO). Based on these intense surveys, the evaluation team offers the following conclusions:

- The overall long-term expected MAR is 0.65. Electric energy efficiency projects had an overall MAR of 0.67, non-electric energy efficiency projects had an overall MAR of 0.42 while on-site generation projects had a MAR of 0.72.
- The MAR for the energy efficiency studies shows a gradually increasing adoption rate flattening after the 5th year from the delivery of the study report to the customer. The MAR for the cogeneration studies indicates the expected delay that results from developing these large projects from the concept stage to the final installation. Overall 95% of the savings are installed within 5 years of the study.
- Of all the measure types, the control measures were observed to have the greatest adoption rate peaking at 0.86 with the lowest adoption rate of 0.20 going to envelope measures.
- The downstate region has a steeper rising MAR curve during the first year compared to the upstate region. On average, the downstate region has a slightly higher adoption rate than the upstate region. Evaluators speculate that the higher energy rates[49] downstate influence customers to take quicker action.
- On-site inspections revealed the need for a net downward correction to telephone interview-reported MARs. Twenty-three out of the fifty sites selected for the on-site surveys had one or more measures that required a MAR correction. The net result was a 7% drop in the MAR value for these studies.
- The electric kWh SRR including large CHP is 0.92 and excluding the large CHP is 0.86. The natural gas SRR excluding the outliers is 0.81. Evaluators also identified net positive changes in savings associated with wood, purchased steam, coal, and fuel oil but realization rates were not calculated. The three most common reasons that evaluators identified for deviations of realization rates from the study predictions were (1) operating differences such as greater or lesser hours of operation per year (35% of measures with "significant" deviation), (2) differences in calculation method (19%), and (3) different analysis parameters.

4.2 PROGRAM RECOMMENDATIONS

The principal goal of the assessment was to analyze the energy savings associated with FlexTech studies completed in 2007-2009. The evaluation team also observed opportunities to improve operation and savings estimation in the future to hopefully increase and narrow the variation in realization rates. Key recommendations include the following:

[49] Based on NYISO (New York Independent System Operator) electric data from 2000 through 2010, the load weighted average wholesale electricity prices for the downstate region are 23% higher than the upstate region.

4.2.1 Program Recommendations

- **Focused marketing of controls studies/vendors** – The MAR survey results indicated that controls measures had the highest adoption rate of all technology groups. Aggressive promotion of this particular type of study could increase the overall program cost-effectiveness by increasing the MAR. Additional investigation regarding why controls measures have such a high adoption rate (is it just due to lower implementation costs and payback times?) could lead to lessons learned, which could be applied to studies associated with other technologies.
- **FlexTech study data post-processing** - Consider adding a step to the final NYSERDA database (buildings portal) data entry. A third-party contractor or the service provider could be hired to review the FlexTech study findings and translate its results into data needed for reporting requirements and impact evaluations. In many cases no translation will be necessary, but for some it will help to:
  - Consistently report the peak demand savings. NYSERDA reports the average demand savings during summer weekday afternoons to the New York Department of Public Service (NY DPS). NYSERDA's tracking system should retain demand savings estimates on the same basis. The FlexTech studies variously reported demand savings as the average demand savings over 12 months, the sum of demand savings for each of 12 months, the annual super–peak demand savings, zero (omitted entirely when it existed), etc. Hence the evaluators found a considerable variability in the reported realization rates associated with the peak electrical demand savings.
  - Capture interactive savings associated with central cooling and heating plants. Some studies focused on energy savings in individual buildings within a large campus, where the buildings were served by central boiler, chiller, or CHP plants. The studies reported savings in Btus of steam or hot water, ton-hours of chilled water, or campus distributed electricity to the building according to the interests of the recipient. In some cases the sub-metered savings were omitted entirely from NYSERDA's tracking system (ton-hours), understating savings, and in other cases the tracking system noted the sub-metered energy impact instead of the impact at the customer's utility meter, potentially distorting NYSERDA's eventual benefit-cost and emissions calculations. NYSERDA should track the net savings at the customer's utility meter, not at the building Btu sub-meters, even if the sub-meter is the parameter of interest to the study recipient. The evaluators recommend that either the service providers be asked to provide such results or that NYSERDA calculate them. This will be more accurate and in at least two cases would have increased the Program's gross reported savings.
- **Add fields in the buildings portal to allow entry of more fuel types** - The SBC Operating Plan reports only on total non-electric MMBtu impacts, not on fuel type. There is a single field in NYSERDA's tracking system for aggregate MMBtu fossil fuel impact and a corresponding single descriptive field for the associated fossil fuel type. Multiple fuel types can only be recorded in a notes section of the database. Large studies and particularly CHP studies often impact multiple fossil fuel energy sources. Separate fields for all fuel types should be considered.
- **Collect and retain electronic files -** Require that service providers deliver PDFs of completed studies and that they submit the Excel analysis files and building model input files (e.g., eQUEST .INP and .PD2 files) and archive them for at least 7 years. Keep electronic copies of the FlexTech study reports linked in the database for easy access. Early in the study start phase the study slowed down because electronic copies of the FlexTech reports were not available. Since the start of this evaluation, most of the sampled study reports have been scanned and connected to the appropriate study. However, significantly larger quantities of older FlexTech study reports still exist only in hard copy format and a few were never found.

- **Incorporate premise identifiers –** This evaluation collected data to supplement program participation overlap analysis not in the scope of this research. The exercise would have been more effective and useful to later researchers and the data would have been easier for them to use if each FlexTech study was associated with a unique premise identifier. A premise ID is a number that is associated with a physical customer address and remains constant over time regardless of occupancy. It is common to and transcends all programs and studies. Premise IDs are commonly used in the utility company environment and in fact it may be possible to collect and use the New York utility companies' premise IDs.[50] Such an identifier also would aid in tracking for marketing and survey fatigue purposes.

4.2.2 Evaluation Recommendations

- Select studies no older than 5 years in the next evaluation cycle – The MAR survey results indicate that measure adoption rates plateau after 5 years of delivering the study to the customer. Hence, the evaluators recommend using studies not older than 5 years from the study date in the next round of evaluations. This will also help end use customers respond more effectively as they are likely to remember the details associated with the report.
- Investigate and develop a more reliable method for the estimation of outside spillover. The outside spillover rate is derived in this evaluation using the same method and survey questions as those in past evaluations. The final outside spillover estimate ends up being based upon a small number of respondents (after dropping those that report no outside spillover). It is a substantial estimate with uncertainty in many of its components. Further research is needed to develop a more reliable method that includes validity checks and is able to better estimate the full impact of the participating service providers on the market.

[50] For IOUs, typically each revenue meter has a unique meter number. One or multiple meters are associated with a customer account number. The utility assigns a permanent Premise ID as part of the customer account number (or tracks it in parallel). When one customer moves out and another in or gets bought, the premise ID part of the customer account number does not change.

Appendix A
Participating Customer Measure Adoption Rate Telephone Questionnaire

NYSERDA FlexTech Impact Evaluation 2009/2010

MAR Telephone Survey

Project ID:	Participant Name:
Participant Contact:	Contact Phone:
Contact Title:	
Participant Contact Address:	
Contact E-Mail:	
Alternative Contact:	Alternative Contact Phone:
Alternative Contact Title:	Alternative Contact E-Mail:
Date FlexTech Report Sent to Customer or Report Approved (RPAP Date):	Total Cost of FlexTech Study:
NYSERDA Project Manager:	NYSERDA Share:
NYSERDA Project Manager Phone:	NYSERDA Project Manager E-Mail:
FlexTech Consultant Company:	Consultant Contact:
Consultant Phone:	Consultant E-Mail:
Date of Survey:	
Interviewer:	

Hello, my name is _____ and I work for ERS. We are under contract to the New York State Energy Research and Development Authority (also known as NYSERDA) and I'm calling on their behalf.

We're contacting organizations that received an energy efficiency study with NYSERDA's assistance under the Flexible Technical Assistance Program (known as FlexTech) from 2003 to 2009. Survey participation is important to NYSERDA and your firm is one of only a sample of projects that have been selected for this evaluation. The input from facility managers, like you, is crucial to our efforts to estimate the program's effectiveness. Without your help, we cannot advance our programs to help more New Yorkers achieve energy efficiency and energy independence. This information will be used to assess program accomplishments and to improve NYSERDA efficiency programs that serve commercial and industrial buildings.

NYSERDA FlexTech Impact Evaluation 2009/2010
MAR Telephone Survey

I have you listed as the contact for the project at [**Put in company name and location**] that participated in NYSERDA's FlexTech Program in [**Enter month and year study completed (RPAP from database)**].

[If they express hesitation, use an appropriate combination of the following.]

Security. Your responses will not affect your ability to participate in the program in the future.

Sales concern. I am not selling anything. I simply want to understand what factors were important to your company when deciding whether to implement recommendations you received as a result of participating in this program.

Contact. If you would like to talk with someone from NYSERDA about this effort, you can call

Evaluation Project Manager: Rebecca Reed, 866-697-3732 Ext. 3559

Obtaining Appropriate Interviewees and Contact Information for MAR Survey v. NTG Survey

We are conducting two types of surveys. The first will be with the Facility Manager or equipment manager that can address technical aspects of whether study recommendations were installed or adapted and then installed. The second will be with a senior decision-maker who decides whether to conduct these types of studies and to approve spending for energy-using equipment or systems.

SCR1: Are you the appropriate person to discuss whether study recommendations were installed?

1. Yes
2. No

[IF SCR1 = YES]:

This survey will take about 20 minutes to complete. We recognize that this is a sizeable time commitment and we can proceed now over the phone; or we can schedule a more convenient time. As another option, we can e-mail or FAX the survey to you to fill out over the next week.

1. Can discuss now
2. Call back on [ ] at (*time*): [ ]
3. Fax survey to: [ ]

[IF SCR1 = NO]:

Could you please direct me to, or provide me with the name of the person who is the most qualified to discuss this project and his/her phone number and e-mail?

(Qualified contact's name) [ ] (Qualified contact's phone) [ ]

(Qualified contact's e-mail) [ ]

NYSERDA FlexTech Impact Evaluation 2009/2010

MAR Telephone Survey

SCR2: Are you the appropriate senior decision-maker to discuss the decision-making process and considerations for conducting the FlexTech study and the decisions to adopt study recommendations, make purchases, installations, or modifications to equipment, systems, or practices?

1. Yes
2. No

[IF SCR2 = YES]:

The decision-maker survey will be conducted in 2-3 months by another NYSERDA evaluation firm, APPRISE.

[IF SCR2 = NO]:

Could you please direct me to, or provide me with the name of the person who is the most qualified to discuss these decisions and his/her phone number and e-mail?

(Qualified contact's name) [ ] (Qualified contact's phone) [ ]

(Qualified contact's e-mail) [ ]

General Instructions

The judgment of program participants is important to this research effort. We are looking for best estimates and opinions as a starting point for characterizing participating projects and assessing the accomplishments of the FlexTech Program.

If you have an opinion or even a rough judgment regarding the answer to a question, please provide it as an estimate for the question.

[If they express hesitation, use an appropriate combination of the following.]

Security. Your responses will not affect your ability to participate in the program in the future.

Sales concern. I am not selling anything. I simply want to understand what factors were important to your company when deciding whether to implement recommendations you received as a result of participating in this program.

Contact. If you would like to talk with someone from NYSERDA about this effort, you can call

Evaluation Project Manager: Rebecca Reed, 866-697-3732 x3559

[Prior to calling, review program records for the project. Have recommended measures and expected savings available.]

Recommended Measures (from FlexTech Study – include measures labeled "R," "I," and "RME" in databases)

(This table will be populated by the survey engineer using the program data prior to the call)

Description of Measures	Savings				
	kWh	kW	MMBtu	Water	O&M
1					
2					
3					
4					
5					

Additional Notes:

NYSERDA FlexTech Impact Evaluation 2009/2010

MAR Telephone Survey

Questions

1. Have you implemented any of the following measures that were recommended in your study?

Study Measure #	(Q1A) Implemented (Y/N/Partial)?	(Q1B) Partially implemented - % of recommended measure savings?	(Q1C) Approximate installation date	(Q1D) Do you plan to implement in the future (Y/N/Unsure)? *If partially implemented, ask for the unimplemented part of the measure*	*Engineer assesses complete measure rejection or clarifies:* Am I correct in understanding that this measure will not be reconsidered? (Y/N)*
1					
2					
3					
4					
5					

* Y = rejected measure that will not be inquired about in future FlexTech MAR surveys.

Additional Notes:

*Ask the next question **only** if you have information that this was a multiple building project.*

2. How many buildings were involved as a part of this project? Their names and addresses?

3. In addition to the funding you received for the feasibility study, did you receive any financial assistance from NYSERDA for the implementation of the recommended measures?

____ Yes *(Go to Question 4)*

____ No *(Go to Question 5)*

____ Don't Know *(Go to Question 5)*

4. If Yes in (3) then:

Description of Measure(s)	NYSERDA Program That Funded the Measure
1	
2	
3	
4	
5	

NYSERDA Program List Reference:

a. Combined Heat & Power
b. Commercial / Industrial Performance Program
c. Loan Fund
d. New Construction Program
e. Existing Facilities Program
f. Peak Load Management Program
g. Multifamily Building Performance Program
h. New York ENERGY STAR® Products Program
i. Solar Electric Incentive Program
j. Business Partners Commercial Lighting Program
k. Wind Incentives for Eligible Installers
l. Renewable Portfolio Standard Customer-Sited Tier Fuel Cell Program
m. Renewable, Clean Energy, and Energy Efficiency Product Manufacturing Incentive Program
n. Clean Energy Business Growth and Development
o. Project Implementation Funding for State Energy Program American Recovery and Reinvestment Act (ARRA)
p. Smart Equipment Choices

(There are more programs than listed here [some expired], so it will be up to the survey engineer to find the appropriate program if the above list does not apply for a given customer.)

5. Did you receive funding or other support for the study or installation from any other sources, such as your utility company, state or federal grants or tax benefits? *Check all that apply. If yes, describe:*

(Q5A) What Received?	(Q5B)What For?	(Q5C) From Whom?
a. Funding	a. Study	a. Utility
b. Technical support	b. Installation	b. State or local govt.
c. Tax benefits	c. Other	c. Federal
d. Other		d. Other

Describe:

_____

_____

_____

_____

6. Did you install any of the other measures that were identified in the study but not recommended? If so, can you tell us why?

7. *If at least one measure has been implemented:* Your site may be selected for an on-site survey which may involve installation of data logging equipment. Data logging equipment will collect data that will be used to determine a project or measure level savings. The measured savings will be compared with the documented estimated savings for that project or measure. Is there a particular time of day or day of the week best to call about scheduling and logistical matters?

8. *If at least part of one measure has not been implemented and is still possibly going to be implemented:* We will contact you again in 9-12 months' time to check on the implementation status of the remaining measures that are recommended but not implemented.

9. We also wanted to let you know that NYSERDA has a Process Evaluation planned for later in 2010 and you may be contacted for a telephone survey for that.

10. Is there anything else you would like to tell us about your participation in the Flex Tech program?

11. Thank you very much for your time!

Appendix B
Sample Design Details

Megdal & Associates

M E M O R A N D U M

To: Judeen Byrne, NYSERDA Energy Analysis and the Evaluation Staff of the New York Department of Public Service (DPS)

From: Kathryn Parlin, WHEC, Satyen Moran, ERS, and Lori Megdal, Megdal & Associates, LLC

Subject: MAR Sampling Design for Flexible Technical Assistance Participants

Date: April 23, 2010

The purpose of this memo is to provide an explanation of the sampling process used to develop the list of projects to include in the Measure Adoption Rate (MAR) telephone survey of NYSERDA's Flexible Technical Assistance (Flex Tech) Program and to discuss the results of that process and the analyses undertaken to verify that the sample is reasonably representative of Flex Tech's activity. The first section describes the overall framework for the analysis. The second and third sections provide a brief summary of the projects and categories and describe the sampling analysis process and the issues that arose.

Sampling Strategy

The Flex Tech program provides cost-sharing to end-use customers for technical assistance studies from energy engineers and experts. When the study is completed, the customer receives a report that identifies cost-effective energy-efficiency measures, operations management, energy procurement and on-site Combined Heat and Power (CHP) opportunities. The Flex Tech program does not require the customer to implement any of the study findings, or to inform NYSERDA if they do. Some participants may obtain further assistance and incentives through other NYSERDA or utility-sponsored programs.

The MAR telephone survey is designed to determine the percentage of the recommended measures and savings that were actually implemented. The on-site survey is designed to derive the savings realization rate using measurement and verification (M&V). Since the on-site survey will have to be conducted at facilities where study findings have been implemented and this information is not available from the program tracking system, the results of the MAR telephone survey will be used to develop the sample frame for the on-site survey. Similarly, the net-to-gross telephone surveys will be conducted for measures that have been implemented.

The design of the sample was influenced by the following factors:

- The results of the MAR telephone survey will be used to construct the sample frame for the on-site survey and for the NTG survey.
- The program recommends measures designed to achieve both electric and natural gas savings, and estimated natural gas savings are substantial in comparison to the electric savings.

- In the most recent FlexTech evaluation completed by Nexant in 2007, the MAR was calculated separately for electric, gas and generation projects. FlexTech reporting has been utilizing this method and FlexTech project staff would like to continue to use this approach.

Consequently, the sampling unit and strategy were designed to provide a suitable sample frame given all of these factors. The remainder of this section covers the sampling method and precision, primary sampling unit, definition of project size, stratification variables and guidelines for sampling.

Sampling Method and Precision

Stratified ratio estimation (SRE) was considered as a potential sampling method since it allows for efficient sampling design and generally requires a lower sample size for a targeted level of precision if there is a strong correlation between the tracking system savings and the implemented savings. However, a review of the previous evaluation done by Nexant indicated that the correlation is not sufficient to create an advantage for the SRE approach, and stratified random sampling was chosen as the sampling method.

The precision/confidence target of 90/10 for the Flex Tech program as a whole (statewide) is specified in the evaluation plan. The MAR sample size of 300 should allow a 90/10 precision/confidence level to be met regionally (upstate/downstate) and also separately for statewide electric savings and gas savings. Sample sizes were verified to be sufficient using a coefficient of variation of 0.50. Data from previous evaluations suggest that the coefficient of variation will be within the range of 0.25 and 0.44.

Primary Sampling Unit

The first step in the sampling process was to conduct a review of the program-level data. For evaluation purposes, it is simpler to consider each site separately. Projects may include recommendations made for more than one site (particularly school districts, municipalities and national chain stores). However, there does not seem to be a reliable way to identify sites consistently from the FlexTech data set, and consequently the sampling was done at the project level. Through the review of the project reports and the MAR survey, the sites associated with each project and the savings at each site will be identified to allow sampling by site for the on-site survey.

There were a total of 1,176 projects with completed reports from program inception through December of 2009. The primary sampling unit is the project, and all recommended measures will be reviewed for the selected projects. This approach allows the evaluators to analyze the project comprehensively and consider interactive effects between measures or groups of measures.

Definition of Project Size

The size of the project in terms of energy savings tends to be a major contributor to the variability of the realization rate. Large projects tend to be more complex and also account for a high percentage of the total program savings. NYSERDA has been clear that both gas and electric savings are central to the underlying goals of the program. Given that some projects have substantial electric savings and little or no gas savings and other projects are the opposite, establishing a method to define the size of the projects was not a trivial task.

After much consideration, the electric savings were converted to source MMBtu and the higher of the two MMBtu values (electric or gas) was selected as the primary variable to define the strata by project size. In combination with other stratification variables discussed below, this approach will allow us to develop a single sample that will be balanced in terms of both gas and electric savings. (Details are provided in the following section.) The sample is designed to estimated the MAR for electric and gas separately to the 90/10 confidence/precision standard.

To convert to source MMBtu, the kWh savings for the electric measures was adjusted to account for savings at the source of generation. This approach avoids the potential pitfall of ending up with a sample that contains a disproportionate number of gas projects, puts the two fuels on a more equal footing in

terms of their actual emissions and accounts for the higher costs to the participant of using electricity. The factor for converting from kWh saved to source MMBtu saved was provided by NYSERDA.[51]

Sampling on maximum source MMBtu tremendously simplifies the sampling process and allows us to use the sample efficiently, *i.e.*, all of the MAR sample with adoptions will be part of the sample frame for the savings realization rate (SRR) site visits. (In contrast, selecting separate gas and electric MAR samples would require that we choose between electric and gas measures for the on-site survey.) This approach will also allow us to be able to obtain the 90/10 confidence/precision target for the SRR with the original sample sizes included in the evaluation plan and used as the basis for the budget. The final evaluation report will include a review of this sampling process and recommendations for future evaluations.

Stratification Variables

The Flex Tech program includes projects that are located both in the upstate and downstate regions. Since the MAR survey has a relatively large sample size (300) and it is the most significant contributor to the overall savings, the sample frame will be stratified by upstate/downstate region for estimating the MAR. However, the program realization rate (RR) will be based on both the MAR and also expensive on-site analyses (which produce the SRR). It would be cost-prohibitive to increase the sample size for the on-site survey to estimate the SRR separately by region and the SRR sample will not be stratified by region.

The timing of the installation is another critical factor to consider. It relates to the lag between when the report is received and when the measures are installed. The previous MAR study showed installations continuing six years or more after the report was provided. Thus, the MAR sample will need to reflect projects with both recent and aging reports. The sample will be allocated to the year of the report according to their representation in the population. The sample will not be designed to estimate the MAR by year to the 90/10 confidence/precision standard.

The third stratification variable is the fuel type. Projects were placed into three categories: 1) gas measures only, 2) gas and electric measures, and 3) electric measures only. Again, the sample was distributed among them in proportion to their representation in the population. The final sample sizes (provided below) should be sufficient to estimate the MAR at the statewide level separately for gas and electric.

Guidelines for Sample Design

The specifics of the sampling plan were established after a review of the data and are based on the following guidelines:

- The sampling method is stratified random sampling.
- The primary sampling unit is the project.
- The sample size is designed to estimate the MAR to a target confidence/precision level of 90/10 for the program as a whole, by upstate/downstate regions and by gas and electric at the statewide level.
- Stratification by size is based on the maximum source MMBtu savings, either electric or natural gas.

[51] The source factor of 9,949.2 Btu/kWh was applied. "This number is based on a three year average (2006, 2007, 2008) and includes a line loss factor of 7.2%. The number is based on natural gas only, since natural gas represents the fuel on the margin." From an e-mail from Rebecca Reid at NYSERDA to Kathryn Parlin, West Hill Energy, on March 15, 2010.

- Sample sizes were determined for the upstate and downstate regions based on their contribution to the overall maximum source MMBtu savings. Sample sizes were also allocated to fuel types based on their proportion of savings in the population.
- The sample was be allocated proportionally according to the year the report was provided to the participant. The sample is not designed to develop MAR estimates separately for each year at the 90/10 confidence/precision target.
- A census of the largest projects will be surveyed.
- The MAR will be calculated separately for gas projects, electric projects and generation projects. Since 89% of the savings from the generation projects is in the census stratum and it is highly likely some other generation projects will be selected in the random sample, there is no need for additional stratification.
- The cut-offs for the strata were determined by a review of the program data.
- The sample will be reviewed to verify that projects with savings from all major measure groups are represented in the sample.
- The projects included in the previous evaluations conducted by Nexant in 2004 and 2006 may also be included in this sample; all previously evaluated projects in the census stratum will be included in the sample and those in the random stratum will be in the sample to the extent that they are represented among the randomly selected projects.

Summary of Projects

Projects met eligibility criteria for the MAR sampling frame based on the time since the participant received the Flex Tech report and savings expected upon adoption of report recommendations. Specific reasons that projects were removed from the sample frame include:

- Projects with reports prior to 2003 were considered to be outside the scope of the study. The Impact Evaluation Team expects that it will not be possible to track down the contacts and obtain reliable information about the status of these projects.
- Projects with reports provided after September 30, 2009 were also excluded to provide additional lag time for the implementation of the recommendations. These projects will be included in future MAR samples.
- Some projects had no electric savings and no gas savings.
- Two projects were large cogeneration projects that were previously removed from the Nexant sample as outliers. These two projects were also eliminated from the MAR telephone survey. These projects have been confirmed as completed and verified savings will be included in the final program savings.

Table 1 below summarizes the removal of projects.

Table 1. Attrition of Projects

Status	# of Projects	Combined Source MMBtu Savings	Total kWh Savings[52]	Total Gas MMBtu Savings
In Sample Frame	657	11,658,395	833,059,345	3,370,127
Report Issued After October of 2009	18	292,812	26,723,695	26,933
Report Issued Before 2003	406	7,081,686	470,009,575	2,405,468
No or Negative Electric Savings	93	0	0	0
Outliers	2	1,946,168	195,610,510	0
Totals	1,176	20,979,061	1,525,403,125	5,802,528

The top level of stratification is by upstate/downstate region. The sample size of 300 was allocated roughly in proportion to the savings within each region, as shown in Table 2 below. Downstate projects account for about 23% of the projects and 24% of the maximum source MMBtu savings.

Table 2: Projects and Total Savings by Region

Region	# of Projects	Maximum Source MMBtu Savings (Electric or Gas)	% of Total Max Source MMBtu Savings	Sample Size
Upstate	508	7,653,056	76%	230
Downstate	149	2,352,028	24%	70
Total	657	10,005,084		300

Within each region, projects were divided into size categories. The stratification of the projects by size is described in

[52] The total electric and gas savings in all of the tables represent the "positive" energy savings, i.e., only gas or electric savings are included and additional energy use is excluded. Generation projects, for example, tend to result in large electric savings and also gas extra use. This approach was taken to avoid including both negative and positive numbers in the calculations, which could have unintended consequences for the MAR. If indicated, separate MAR's may be calculated for gas and electric extra use.

below. The cut-off points were established by reviewing the project-level data.

and Table 4 show the number of projects, the maximum source MMBtu savings and minimum and maximum source MMBtu savings by project for each stratum.

The strata of very small projects, consisting of 176 projects that make up a total of 3% of the total program annual energy savings in the upstate region (49 projects sites in the downstate region), were excluded from the sample, since verifying these smaller projects would require resources while not contributing to reducing the uncertainty of the MAR estimates. For the projects in the stratum designated for random sampling, the average project size (in terms of maximum source MMBtu) is somewhat smaller in the upstate region (9,403 annual source MMBtu as compared to 10,753 MMBtu in the downstate region).

Table 3: Project Sites and Savings by Size Stratum for the Upstate Region

Stratum	Sampling Method	# of Projects	Sum of Maximum Source MMBtu Savings (Electric or Gas)	% of Maximum Source MMBtu Savings	Min Source MMBtu	Max Source MMBtu	Target Sample Size
0	None	176	223,370	3%	26	2,948	0
1	Random	262	2,463,639	38%	2,985	24,805	160
2	Census	70	4,966,047	59%	25,244	620,908	70
Total		508	7,653,056		26	620,908	230

Table 4: Projects and Savings by Size Stratum for the Downstate Region

Stratum	Sampling Method	# of Projects	Sum of Maximum Source MMBtu Savings (Electric or Gas)	% of Maximum Source MMBtu Savings	Min Source MMBtu	Max Source MMBtu	Target Sample Size
0	None	49	68,741	3%	263	2,717	0
1	Random	76	817,253	35%	2,841	24,963	46
2	Census	24	1,466,034	62%	26,000	209,393	24
Total		149	2,352,028		263	209,393	70

Table 6Table 5 and Table 6 show the distribution of projects and savings by the year of the report. The sample was allocated to the year proportionally according to the percentage of the source MMBtu savings in each report year.

Table 5: Projects and Distribution of Savings by Year for the Random Sample Stratum in the Upstate Region

Year	# of Projects	% of Projects	% of Maximum Source MMBtu Savings (Electric or Gas)	% of Total MWh Saved	% of Gas MMBtu Saved	Target Sample Size
2003	68	26.0%	24.8%	26.0%	18.8%	42
2004	38	14.5%	13.4%	12.3%	14.6%	23
2005	46	17.6%	19.4%	19.8%	19.9%	28
2006	32	12.2%	11.6%	10.6%	13.6%	20
2007	29	11.1%	13.4%	13.6%	15.0%	18
2008	35	13.4%	13.7%	13.7%	14.2%	21
2009	14	5.3%	3.7%	4.0%	4.0%	9
Total	262					161

Table 6: Projects and Distribution of Savings by Year for the Random Sample Stratum in the Downstate Region

Year	# of Projects	% of Projects	% of Maximum Source MMBtu Savings (Electric or Gas)	% of MWh Saved	% of Gas MMBtu Saved	Target Sample Size
2003	22	28.9%	30.0%	27.1%	37.8%	13
2004	18	23.7%	26.9%	27.5%	27.3%	11
2005	7	9.2%	9.5%	10.1%	9.3%	4
2006	7	9.2%	6.7%	7.2%	0.0%	4
2007	5	6.6%	10.4%	11.2%	3.2%	3
2008	8	10.5%	9.5%	9.8%	18.8%	5
2009	9	11.8%	7.1%	7.0%	3.6%	5
Total	76					45

The final stratification variable is the fuel type of the project. As can be seen in Table 7 below, there are 187 projects with gas savings in the random sample and 333 projects with electric savings. Proportional allocation to gas only, gas/electric and electric only projects results in sample sizes of 118 gas and 205 electric projects in the random sampling stratum. These samples sizes were verified to be sufficiently large to meet the 90/10 confidence/precision standard for electric and gas.

Table 7: Projects and Distribution of Savings by Fuel Type

Stratum	Sampling Method	# of Projects	# of Projects with Gas Only Measures	# of Projects with Gas and Electric Measures	# of Projects with Electric Only Measures	Total MWh Savings	Total Gas MMBtu Savings
0	None	225	7	109	109	25,824	110,948
1	Random	338	5	182	151	289,005	1,110,896
2	Census	94	1	43	50	518,231	2,148,283
Total		657	13	334	310	833,059	3,370,127

Comparison to Previous MAR Studies

Nexant conducted MAR surveys for the Flex Tech program in 2004 and 2006. From the database provided by NYSERDA for these surveys, it appears that 210 projects were included in the Nexant sample and 127 surveys were completed. There are three major differences between the Nexant sample selection and the current evaluation, as described below.

- It appears that the initial MAR sampling was done on the basis of the electric savings and Nexant subsequently separated its sample into generation projects, electric efficiency projects and projects with fossil fuel savings for the MAR analysis. In the current sampling plan, the sampling is based primarily on the size of the projects (as measured in annual source MMBTu savings) and region. In addition, projects were selected by the year of the project report and fuel type proportionally to their representation in the population. This approach should provide separate

MAR estimates by electric and gas savings. The vast majority of the generation savings (89%) is in the census stratum, so it will also be possible to estimate the MAR for generation projects.

- Nexant appears to have assigned projects to bins depending on the time lag between when the report was provided and when the MAR telephone call was made. In the current evaluation, the plan is to assign projects to bins according to the time difference between when the report was provided and when the measure was installed, where that information can be reliably obtained.
- For the fossil fuel MMBtu's, Nexant apparently combined savings and extra use into one field and estimated the MAR on this basis. The Impact Evaluation Team is planning to separate savings from extra use for both natural gas and electricity. This approach ensures that the results of the MAR survey and site visits will produce appropriate and realistic adjustments to the claimed savings. The results of the MAR survey will determine whether it is necessary to estimate MAR values for extra energy use.

A number of analyses were conducted to assess the implications of the changes to the approach between the previous and current evaluations. The first step was to compare the projects included in the previous Nexant evaluations to the current sample frame. The projects were identified that were in our initial sample frame and also within Nexant's sample. It appears that Nexant completed surveys for 72 projects in the current sample frame.[53] The Nexant sample predominantly consists of projects with electric savings, and in order to be able to estimate the MAR by gas and electric, the current sample will need to be distributed proportionally among projects with gas only, electric/gas and electric only savings. Table 8 below shows the number of projects that were included in Nexant's survey.

Table 8: Comparison of Sampling Frame with Previous MAR Surveys

Stratum	Sampling Method	Total # of Projects	# of Projects in the Nexant 2004/2006 Sample	# of Projects with Completed MAR Survey in 2004/2006 Studies	# of Projects Fully Implemented in 2004/2006 Studies
Downstate					
0	None	49	0	0	0
1	Random	76	19	12	1
2	Census	24	4	4	1
Upstate					
0	None	176	4	2	0
1	Random	262	47	37	4
2	Census	70	28	17	1

Our initial plan had been to incorporate the projects that completed the Nexant MAR survey into our analysis. However, the substantial differences in the sampling strategies make it impractical to do so. Rather, the sample frame will include all eligible projects sites (including those previously surveyed by Nexant), and the sample will be randomly selected within the strata. The sample will be reviewed to identify the projects sites from the prior MAR survey and those randomly selected projects from the

[53] Some of the 115 projects in Nexant's survey were removed from the sample frame in the data cleaning process. Forty-two (42) had reports provided prior to 2003 and two (2) were removed by NYSERDA because NYSERDA did not contribute to the costs of the TA study. In addition, twelve (12) projects with completed surveys by Nexant were not found in the current FlexTech program database.

earlier studies will be used in the MAR analysis. Any selected projects sites that were found to have installed all of the recommended measures in the 2004 or 2006 MAR studies will not be contacted again, and will be included in the analysis with the full savings for the purpose of calculating the MAR. As can be seen in Table 8, 19 of the 94 census projects were in the sample from the previous evaluations and will be included in the current MAR study.

The projects were also reviewed to assess whether it was necessary to stratify by generation projects. Nexant's approach to establishing a separate MAR for generation projects was based on the assumption that these projects have a longer lead time, and, thus, are much slower to reach completion. While this assumption appears to be accurate, a thorough review of the current data led us to the conclusion that it is not necessary to draw separate samples by generation versus energy efficiency. The majority of the generation projects are in the census stratum and will thus not have a disproportionate impact on the overall realization. As can be seen in Table 9 below, generation projects account for only 7% of the savings in the random stratum for upstate and 10% for downstate.[54] About 89% of the MWh savings associated with generation projects is included in the census stratum.

Table 9. Impacts of Generation Projects

Stratum	# of Projects	Total MWh Savings	# of Projects with Generation	MWh Savings from Generation Projects	% of Projects with Generation	% of MWh Savings from Generation
Upstate						
0	176	19,400	5	688	3%	4%
1	262	212,923	12	15,003	5%	7%
2 (census)	70	386,306	11	127,386	16%	33%
Total	508	618,629	28	143,077		
Downstate						
0	49	6,425	0	0	0%	0%
1	76	76,081	8	7,733	11%	10%
2 (census)	24	131,925	8	61,505	33%	47%
Totals	149	214,431	16	69,238		

To the extent that projects from the previous Nexant evaluations are selected for the sample, the difference in the definition of the timing bins will introduce some complexity to the analysis. t appears that the actual installation dates for the projects in the Nexant surveys are not included in the data set provided and may necessitate making assumptions regarding the timing of the installation based on other responses to the previous MAR survey. For those participants included in the current sample, additional questions may be added to the survey. The Impact Evaluation Team will address these issues through the survey instrument development and analysis of results.

Conclusion

The primary sampling unit for the MAR survey is the project. The sample is designed to estimate the MAR to meet the 90/10 confidence/precision target for the entire program, for the upstate/downstate regions and by fuel type (gas and electric) at the statewide level. The sample was stratified by projects

[54] Two extremely large generation projects have been analyzed separately as outliers.

size (as defined by the maximum electric or gas source MMBtu savings) and region. The sample was allocated proportionally to the year of the report and the fuel type (electric or gas) based on the number of projects. The sample size for each fuel type was verified to ensure that the 90/10 standard can be met based on the variation found in the previous Nexant evaluation. Projects included in Nexant's 2004 and 2006 MAR surveys will be included in the sample frame; the 19 previously-evaluated projects in the census stratum will be included in the current MAR survey, and projects in the random stratum will be chosen if they are selected to be in the sample.

The final sample size for the Flex Tech MAR survey to be conducted as part of the Flex Tech impact evaluation is 300 projects, consisting of 230 in the upstate region and 70 downstate. For the upstate region, 70 were selected in the census stratum and 160 in the randomly selected stratum. In the downstate region, the total sample size of 75 consists of 24 projects sites in the census stratum and 46 in the random stratum.

Flex Tech impact evaluation Champion: Satyen Moray, Energy & Resource Solutions, Inc. (ERS)

Flex Tech impact participant sample design and sampling: Kathryn Parlin, West Hill Energy & Computing, Inc.

Appendix C
On-Site Visit Checklist & Inspection Form Regarding Investment Practices

NYSERDA FlexTech On-Site Visit Checklist & Survey

Site Name:

Engineer Name:

Date:

PRELIMINARY SITE INFORMATION

Contact Name & Title: _____

Contact Phone Number: _____

Site Address and Site Visit Details: _Location to Meet On-Site:

_____

☐ **If loggers will be deployed, then verify that an electrician would be available during the site visit**

☐ **If a combustion test will be performed inquire if the boiler exhaust stack has a port to insert the combustion analyzer probe**

☐ **Check the weather a day or two before the scheduled site visit and call to confirm the time and location one more time**

PRE-SITE VISIT CHECKLIST

Equipment to bring on-site (depending on your M&V plan):

- ☐ **Site visit kit** - check that it includes a flashlight, multi-tool, screw driver, gloves, glasses, tape measure, and zip ties
- ☐ **Flow meter and transducers** – check that it includes both small and large pipe transducers
- ☐ **Pipe size guide** - guide should include steel, copper, and PVC standard sizing information
- ☐ **Drill bit for penetrating boiler exhaust piping** (piping is likely to be stainless steel or PVC)
- ☐ **Plugs or high temperature tape** – plugs or high temperature tape capable of withstanding the exhaust stack temperature to cover hole made to test combustion efficiency
- ☐ **Pipe tape** - pipe tape to re-tape insulation around pipes after flow measurements
- ☐ **Steel wire** - to affix sensors to supply and return water pipe
- ☐ **Pipe insulation** – to insulate temperature sensors against supply and return water pipes
- ☐ **Temperature sensor loggers**
- ☐ **Temperature sensors**
- ☐ **Thermocouples and thermocouple loggers**
- ☐ **Motor on/off loggers or CTS and loggers**
- ☐ **Logger communication cable**
- ☐ **Computer for launching loggers**
- ☐ **Directions to the site**
- ☐ **Combustion analyzer**
- ☐ **Camera**

SITE INSTALLATION SURVEY

During the site walk-through observe the general condition of and types of the various energy systems (HVAC, lighting, motors, chillers, compressors, pipe insulation, building envelope, etc.) in the facility. Particular focus should be on candidates for the same types of efficiency measures as you are studying. The intent of this survey is to cost-effectively note whether the firm appears to be willing to make investments to lower long-term operating costs (such as making investments for energy efficiency) versus showing a policy of lowest immediate costs. The minimizing of current costs may be easily observed without further inquiry. Is there a lot of old inefficient equipment in use? Does the facility appear to "jerry-rig or duct tape" things together when a much smarter investment would be more cost-effective for the firm, even over just the medium-term?

If high efficiency equipment beyond what is part of the project under study is observed then conversation may be needed to separate out a firm's equipment investment policy from impacts from efficiency programs. This would include (1) identifying replacement patterns beyond the evaluated measures then, (2) inquire as to whether the observed non-evaluated projects were implemented with or without NYSERDA or other efficiency funding (such as rebates/grants from other organizations, tax credits); and (3) if their installation preceded or followed or were inspired by the measure installation through the NYSERDA program.

As an example, if the FlexTech study had recommended retrofitting T12 lighting with T8 lighting in the office, then your role while at the site is to observe if other areas outside of the office also feature T8 or better lighting technology. If you find that the warehouse or production areas are using T8s or other similar higher efficiency lighting system then you should ask your site contact if those lights received NYSERDA or other funding to support the efficiency investment. If no, also ask if the warehouse or production area T8 retrofits decisions were made based on the office T8 project results, or if the decisions were independent of the evaluated project.

We want this effort to be seamless with the walk-through. We want no significant additional time required from you or the participant. This effort is to collect information as is readily available/observable at the same time as the planned M&V effort. Sometimes the firm's policies in this area might be easily determined. If not, just note what you can from your observations.

The following subsections should be used to guide the data collection efforts.

1. **FOR SRR CENSUS PROJECTS ONLY** (IN THE ON-SITE SAMPLE, THIS IS DIFFERENT FROM THE MAR SAMPLE)
    a. Does your organization have any written policies about buying energy efficient equipment or otherwise implementing energy efficiency projects? Can you provide a copy?

    b. Search the websites of the company you are visiting for corporate sustainability statements and similar efficiency-related announcements or policies.

MEASURE RELATED DATA COLLECTION TABLE

Measure Description	**Does all observed of this equipment type appear to be program measures? (Y/N)**	
	Yes No	If Yes, then do NOT include equipment type in table below.
	Yes No	
	Yes No	
	Yes No	
	Yes No	
	Yes No	

Condition/Efficiency Rating Scores

0	"Duct-taped" to make barely operational. Very cost-effective to fix correctly or replace. (Unbelievable, looks like they are ready to go out of business.)
1	Old and obvious that replacement should have been undertaken to lower operating/repair costs.
2	Operational but retrofit, for energy efficiency or repair costs, quite likely to be cost-effective.
3	Not high efficiency but not old and meets standard industry practice.
4	High efficiency equipment as of a few years ago. Payback for retrofit probably at least several years out.
5	High efficiency equipment in good operating condition.

Equipment Type Observed*	Condition/Efficiency Rating (Proportion of Equipment that is….)			For those with Rating of 4 or 5				
	0, 1 or 2	3	4 or 5	Inquiry Made (Y/N)	Installed with NYSERDA funding	Installed with other efficiency funding	Part of firm's standard practice	Unknown

* Beyond pieces of equipment included in the program project being evaluated.

Example Tables

Measure Description	**Does all observed of this equipment type appear to be program measures? (Y/N)**	
0.5 kW/ton Chiller	X Yes No	If Yes, then do NOT include equipment type in table below.
AHU VFDs	X Yes No	

Equipment Type Observed*	Condition/Efficiency Rating (Proportion of Equipment that is….)			For those with Rating of 4 or 5				
	0, 1 or 2	3	4 or 5	Inquiry Made (Y/N)	Installed with NYSERDA funding	Installed with other efficiency funding	Part of firm's standard practice	Unknown
Warehouse lighting	100%	0	0					
Office lighting		25%	75%	Y	√			
Boiler			100%	Y		√		
Air leakage at fenestration	75%		25%	Y				√
Pipe insulation		100 %						

Appendix D
Participating Customer Net-To-Gross Telephone Questionnaire

NYSERDA Flexible Technical Assistance (FlexTech)
Program Impact Evaluation 2010 Participating Owner Survey
FINAL Survey Instrument
05/05/2011

[IF INTRO=01, READ INTRO1, ELSE SKIP TO INTRO2]
INTRO 1

Hello may I please speak to **[NAME_MAR]**? I'm calling on behalf of the New York State Energy Research and Development Authority (or NYSERDA).

In our survey on equipment adoptions conducted a couple of months ago regarding **[ACCOUNT_SITE]**'s participation in NYSERDA's FlexTech Program, we had indicated that there could be an additional two surveys and a site visit to complete NYSERDA's evaluation research. This is the decision maker survey that was mentioned. At the end of that survey, you had indicated that we could contact you for our decision maker survey. **[GO TO Q1]**

[IF INTRO=02, READ INTRO2, ELSE SKIP TO INTRO3]
INTRO 2

Hello may I please speak to **[NAME_MAR]**? I'm calling on behalf of the New York State Energy Research and Development Authority (or NYSERDA).

In our survey on equipment adoptions conducted a couple of months ago regarding **[ACCOUNT_SITE]**'s participation in NYSERDA's FlexTech Program, we had indicated that there could be an additional two surveys and a site visit to complete NYSERDA's evaluation research. This is the decision-maker survey that was mentioned. **[GO TO Q1]**

[IF INTRO=03, READ INTRO3, ELSE SKIP TO INTRO4]
INTRO 3

Hello may I please speak to **[NAME_DM]**? I'm calling on behalf of the New York State Energy Research and Development Authority (or NYSERDA).

Your firm was selected as part of a small carefully designed sample. A couple of months ago we interviewed **[NAME_MAR]** regarding equipment adoption relating to the project at **[ACCOUNT_SITE]** that participated in NYSERDA's Flexible Technical Assistance Program, also known as FlexTech**. [NAME_MAR]** recommended that you might be the best person to speak with regarding the decision-making for this project or projects similar to this one. **[GO TO Q1]**

[IF INTRO=04, READ INTRO4, ELSE SKIP TO INTRO5]

Hello, may I please speak to [NAME_DM]? I'm calling on behalf of the New York State Energy Research and Development Authority (or NYSERDA).

We are conducting research on NYSERDA's Flexible Technical Assistance Program (also known as FlexTech) and would like to ask you some questions regarding your organization's decision to participate in FlexTech. Our records indicate that you might be the best person to speak with regarding the decision making for this project or projects similar to this one. [GO TO Q1]

[IF INTRO=05, READ INTRO5]

INTRO 5

Hello may I please speak to **[NAME_DM]**? I'm calling on behalf of the New York State Energy Research and Development Authority (or NYSERDA).

We are conducting research on NYSERDA's Flexible Technical Assistance Program, also known as FlexTech. We recently conducted a survey with **[COMP_DM]**, and he/she indicated that you were influential in the decision making for this project or projects similar to this one. **[GO TO Q1]**

Q1. Are you the appropriate person to discuss issues related to the decision-making at your firm to undertake the FlexTech study and then to decide to adopt at least some of its recommendations?

01 YES
02 NO, NOT CORRECT RESPONDENT
96 REFUSED
97 DON'T KNOW

[IF Q1=01, SKIP TO SCR-1, ELSE ASK Q2]

Q2. Can you provide me with a contact name and phone number for a person in your organization that was involved in that decision-making?

01	YES	[**SPECIFY NAME, NUMBER**]
02	NO	**[TERMINATE]**
96	REFUSED	**[TERMINATE]**
97	DON'T KNOW	**[TERMINATE]**

[IF RECOMMENDED DECISION-MAKER CONTACT RECOMMENDS SOMEONE ELSE, RECORD NEW CONTACT NAME, TITLE, AND PHONE NUMBER]

[CONTACT THIS PERSON, REPEAT INITIAL INTRODUCTION (INTRO 4) AND CONTINUE WITH THE FOLLOWING QUESTION]

[SCHEDULING]

SCR-1. This survey will take about 25 minutes **[CORRECT BASED ON TIMING/ TESTING]** to complete. Can we discuss the project now, or can we schedule a time when I can call you back?

1. CAN DISCUSS NOW **[PROCEED TO SECTION DM: PROJECT SPECIFIC REVIEW]**
2. CALL BACK ON:_____ AT TIME:_____

[READ IF NECESSARY: Your feedback is very important to this research and can enable NYSERDA to improve its program for firms such as yours. Your responses to this survey will be kept confidential to the extent permitted by law.]

[IF CENSUS==01, ASK DM1–DM3, ELSE SKIP TO REP1.]

SECTION: DEM – ADDITIONAL DECISION-MAKERS

DM1a. Generally, how are decisions related to equipment and the facility made at your firm? I'm going to read a list of decision descriptions, and I would like to know which statement **best** describes how each decision is made at your firm. [READ LIST. MARK ONE RESPONSE ONLY.]

01 A committee which I chair has final say in the decision.
02 The decision is completely a committee decision.
03 Someone else makes the technical recommendations but I have the final financial or contracting authority.
04 I make the recommendations but others have the financial or contracting authority.
05 I make recommendations and the corporate office elsewhere makes the decision, but my recommendations are normally followed.
06 I make recommendations but the corporate office always makes their own decisions, sometimes with little regard to my recommendations.
07 There are multiple groups and decision points that must be passed that are more complicated than these other statements.
96 REFUSED
97 DON'T KNOW

[IF DM1a=01-06, SKIP TO DM2. IF DM1=07, 96, OR 97, ASK DM1b]

DM1b. Please describe the decision-making process.

01 [RECORD VERBATIM]
96 REFUSED
97 DON'T KNOW

DM2. Who played key roles in the decision-making process?

[FIRST, HOW MANY PEOPLE PLAYED KEY ROLES?
CAN YOU GIVE ME THEIR NAMES AND TITLES?
ASK DM3.
THEN, OBTAIN PHONE NUMBERS, EMAIL ADDRESSES]

[ENSURE YOU HAVE ALL THE PEOPLE THAT CORRESPOND TO THE RESPONSE IN DM1. INQUIRE:
--WHO IS ON THE COMMITTEE FOR COMMITTEE DECISIONS
--WHO IN THE CORPORATE OFFICE IF THEY HAVE INPUT INTO THE DECISIONS
--WHO IS/ARE THE FINANCIAL AND CONTRACTING AUTHORITIES IF THEY ARE INVOLVED.]

[ACCEPT UP TO 5 ENTRIES]

Title	Name	Phone	Email	DM3 Score

DM3. On the scale of 0 to 4, with 0 being no influence and 4 being very influential, how influential was each person in the decision making process? **[RECITE THE NAME(S) OBTAINED IN THE PREVIOUS QUESTION.]**

00 NO INFLUENCE
01
02
03
04 VERY INFLUENTIAL

[IF DM3=03 OR DM3=04, WE MAY CONDUCT ADDITIONAL SURVEY WITH THE NEW CONTACT.]

SECTION: REP – EARLY REPLACEMENT VERSUS RETROFIT

Now, I'd like to ask you some questions about the specific measures you installed at **[ACCOUNT SITE]**. I'd like you to focus on the following measures – **[LIST REP_A-REP_E]**

[IF DIFFERENT MEASURES AT DIFFERENT SITES, LIST AND ASK ABOUT ALL MEASURES AT ONE SITE, AND THEN GO TO THE NEXT SITE AND LIST MEASURES]

[FOR EACH INSTALLED MEASURE FROM SAMPLE FILE, ASK REP1-REP2 IN SEQUENCE THEN GO TO NEXT MEASURE AND ASK REP1-REP2.]

REP1. To the best of your recollection, how old was the equipment you replaced with **[REP_A-REP_E]**? [READ IF NECESSARY: Do you know what year it was originally installed?]

[FOR EACH APPLICABLE MEASURE, RECORD AGE IN YEARS OR ORIGINAL INSTALL DATE]

Prior Equipment for Up to 5 Measures with the Highest Installed Savings	Age (Years) -	OR	Original Install Date	Reasons for Replacement
[REP_A]	REP1_a1		REP1_a2	REP2_a
[REP_B]	REP1_b1		REP1_b2	REP2_b
[REP_C]	REP1_c1		REP1_c2	REP2_c
[REP_D]	REP1_d1		REP1_d2	REP2_d
[REP_E]	REP1_e1		REP1_e2	REP2_e

REP1_a-e

01 [RECORD AGE (YEARS)]
02 [RECORD DATE (YEAR)]
03 DID NOT REPLACE OLDER EQUIPMENT
94 NOT YET INSTALLED
96 REFUSED
97 DON'T KNOW

[IF REP1_a-e=3, SKIP TO REP2b_a-e, ELSE READ REP2_a-e]
REP2_a-e. Which of the following **best** describes your decision to replace **[REP_A-REP_E]**? [READ LIST. MARK ONE RESPONSE ONLY.]

01 It was working but not as efficiently as newer models
02 It was working but we needed a larger/smaller system
03 It was working but old and would probably need to be replaced in next couple of years anyway
04 It required frequent maintenance
05 It was not working
06 Other (Specify)
94 NOT YET INSTALLED
96 REFUSED
97 DON'T KNOW

[IF REP1_a-e=3, ASK REP2B_a-e. ELSE SKIP TO AUD1]
REP2b_a-e. Why did you decide to incorporate [REP_A-E]?

01 [RECORD VERBATIM]
96 REFUSED
97 DON'T KNOW

SECTION: AUD – INFLUENCE OF AUDIT

AUD1. If the FlexTech Program or its assistance had not been available to you, what is the likelihood that you would have performed some type of study, even one of a **lower quality** than the study performed under the FlexTech Program? Please rate your likelihood in percentage terms with 0% being not at all likely and 100% being extremely likely.

01 [RECORD PERCENT. ACCEPT [0-100]]
96 REFUSED
97 DON'T KNOW

[IF AUD1>0, ASK AUD2. IF AUD1=0, SKIP TO FR1]

AUD2. If the FlexTech Program or its assistance had not been available to you, what is the likelihood that you would have performed a study of **at least the same or similar quality** to the one performed under the FlexTech Program? Again, please rate your likelihood in percentage terms with 0% being not at all likely and 100% being extremely likely.

[READ IF NECESSARY: In the previous question, I was asking about the likelihood you would have performed **any** type of study. This question asks about the likelihood you would have performed a study of **the same or similar quality** as the FlexTech study.]

01 [RECORD PERCENT. ACCEPT [0-100]]
96 REFUSED
97 DON'T KNOW

AUD3. Did the FlexTech Program or its assistance cause you to conduct the study earlier than you would have without the program?

01 YES
02 NO
96 REFUSED
97 DON'T KNOW

[IF AUD3=01 – YES, ASK AUD4. ELSE, SKIP TO FR1]

AUD4. How much earlier?

01 [RECORD NUMBER OF MONTHS]
02 [RECORD NUMBER OF YEARS]
96 REFUSED
97 DON'T KNOW

SECTION: FR – FREE RIDERSHIP

FR1. Prior to participating in the FlexTech Program, were you planning to install any of the energy efficiency or demand measures recommended by the FlexTech report?

01 YES
02 NO
96 REFUSED
97 DON'T KNOW

[IF FR1=01, ASK FR2, ELSE SKIP TO FR3]

FR2. Could you please describe any plans that you had to incorporate the measures prior to participating in the FlexTech Program?

[RECORD VERBATIM]

[BASED ON RESPONSES TO FR2, FILL IN A "0 TO 4" SCORE INDICATING THE EXTENT TO WHICH RESPONDENT WAS ALREADY PLANNING TO INCORPORATE THE ENERGY EFFICIENCY MEASURE. DO NOT ASK RESPONDENT DIRECTLY. "0" INDICATES THAT RESPONDENT HAD NO PLANS AT ALL; "4" INDICATES THAT RESPONDENT HAD DOCUMENTED PLANS AND HAD BUDGETED FOR ALL OF THE EFFICIENCY MEASURES.]

[NO PLANS] **[DOCUMENTED PLANS/BUDGET]**

0 1 2 3 4

[FULL GUIDELINES FOR ASSIGNING HIGH-EFFICIENCY PROJECT PLANNING SCORE]

SCORE	EXTENT OF PLANNING
0	NO PLANS FOR HIGH EFFICIENCY EQUIPMENT; RESPONDENT MAY HAVE CONSIDERED ALTERNATIVE TECHNOLOGY OPTIONS, BUT DID NOT EXPLICITLY CONSIDER HIGH EFFICIENCY.
1	INITIAL STEPS TOWARD CONSIDERATION OF HIGH EFFICIENCY SUCH AS REQUESTING INFORMATION ON OR DISCUSSING, IN GENERAL, HIGH EFFICIENCY OPTIONS WITH VENDORS OR CONTRACTORS.
2	IN-DEPTH DISCUSSION OR CONSIDERATION OF SPECIFIC TYPES OF HIGH EFFICIENCY EQUIPMENT (E.G., LIGHTING, HVAC, APPLIANCES), INCLUDING THEIR POSITIVE AND NEGATIVE ATTRIBUTES AND COSTS.
3	IDENTIFICATION OF SPECIFIC EQUIPMENT MANUFACTURERS AND MODELS, INCLUDING ASSESSMENT OF THEIR RELATIVE COSTS AND PERFORMANCE CHARACTERISTICS.
4	HIGH EFFICIENCY EQUIPMENT AND DESIGNS FULLY SPECIFIED AND EXPLICITLY SELECTED OR INCORPORATED INTO PROJECT BUDGET.

[SHORT GUIDELINES FOR ASSIGNING HIGH-EFFICIENCY PROJECT PLANNING SCORE]

[PLEASE NOTE: CATI PROGRAM WILL ONLY INCLUDE THE SHORT GUIDE]

SCORE	EXTENT OF PLANNING
0	NO PLANS
1	TOOK INITIAL STEPS DISCUSSED EQUIPMENT OPTIONS GENERALLY WITH CONTRACTORS
2	DISCUSSED EQUIPMENT OPTIONS IN-DEPTH CONSIDERED SPECIFIC TYPES OF EQUIPMENT DISCUSSED POSITIVE AND NEGATIVE ASPECTS
3	IDENTIFIED SPECIFIC EQUIPMENT MANUFACTURERS AND MODELS ASSESSED RELATIVE COST AND PERFORMANCE
4	SPECIFIED HIGH EFFICIENCY EQUIPMENT INCORPORATED EQUIPMENT INTO BUDGET

[SEE HANDOUT TITLED *"FR2"* FOR ADDITIONAL ASSISTANCE WITH THIS QUESTION]

FR3. Thinking about the measures you incorporated, did your participation in the FlexTech Program influence the type or amount of measures you selected or their efficiency levels?

01 YES
02 NO (all equipment would have been incorporated at the same high efficiencies)
96 REFUSED
97 DON'T KNOW

[IF FR3=01 – YES, ASK FR4, ELSE SKIP TO FR5]

FR4. Please briefly describe how the FlexTech Program influenced your decision to incorporate high efficiency measures at this site.

[BASED ON RESPONSE TO FR4, FILL IN A "0 TO 4"SCORE INDICATING THE EXTENT TO WHICH THE PROGRAM INFLUENCED THE DECISION TO INCORPORATE HIGH EFFICIENCY MEASURES. DO NOT ASK RESPONDENT DIRECTLY. "0" INDICATES THAT THE PROGRAM HAD NO INFLUENCE; "4" INDICATES THAT THE PROGRAM WAS THE PRIMARY REASON THAT HIGH EFFICIENCY MEASURES WERE INCORPORATED.]

[NO PROGRAM INFLUENCE] **[PROGRAM PRIMARY INFLUENCE]**

0 1 2 3 4

[FULL GUIDELINES FOR ASSIGNING PROGRAM INFLUENCE SCORE]

SCORE	CHARACTERIZATION OF PROGRAM INFLUENCE
0	NO INFLUENCE ON THE DECISION TO INSTALL HIGH EFFICIENCY EQUIPMENT. ALL EQUIPMENT WOULD HAVE BEEN INSTALLED AT THE SAME EFFICIENCIES EVEN WITHOUT THE PROGRAM.
1	PROGRAM HELPED IN MAKING FINAL DECISION ON EQUIPMENT THAT HAD ALREADY BEEN THOROUGHLY CONSIDERED.
2	PROGRAM LENT CREDIBILITY TO THE DECISION TO INVEST IN HIGH

	EFFICIENCY AND/OR IT PROVIDED INFORMATION THAT HELPED EXPAND THE QUANTITY, SCOPE, OR EFFICIENCY OF THE EQUIPMENT.
3	PROGRAM IDENTIFIED A SIGNIFICANT NUMBER OF SPECIFIC HIGH EFFICIENCY OPTIONS THAT WERE INSTALLED BUT THAT HAD NOT PREVIOUSLY BEEN CONSIDERED AND/OR PROGRAM WAS A MAJOR DRIVER BEHIND A SIGNIFICANT INCREASE IN THE QUANTITY, SCOPE, OR EFFICIENCY OF HIGH EFFICIENCY EQUIPMENT.
4	PROGRAM WAS THE PRIMARY REASON THAT HIGH EFFICIENCY EQUIPMENT WAS INSTALLED IN THE PROJECT.

[SHORT GUIDELINES FOR ASSIGNING PROGRAM INFLUENCE SCORE]

[PLEASE NOTE: CATI PROGRAM WILL ONLY INCLUDE THE SHORT GUIDE.]

SCORE	CHARACTERIZATION OF PROGRAM INFLUENCE
0	NO INFLUENCE
1	HELPED MAKE FINAL DECISION
2	PROGRAM LENT CREDIBILITY TO DECISION PROGRAM PROVIDED INFO TO HELP EXPAND SCOPE OF PROJECT (e.g., QUANTITY, EFFICIENCY)
3	PROGRAM IDENTIFIED OPTIONS THAT HAD NOT PREVIOUSLY BEEN CONSIDERED PROGRAM WAS A MAJOR DRIVER BEHIND INCREASE IN PROJECT SCOPE (.e.g, QUANTITITY, EFFICIENCY)
4	PROGRAM WAS PRIMARY REASON THAT HIGH EFFICIENCY EQUIPMENT WAS INSTALLED IN THE PROJECT.

[SEE HANDOUT TITLED *"FR4"* FOR ADDITIONAL ASSISTANCE WITH THIS QUESTION]

FR5. On a scale of 0 to 4, where 0 = "not at all important" and 4 = "very important," please indicate how important the FlexTech Program was in your decision to incorporate high efficiency measures at this site?

[NOT AT ALL IMPORTANT] **[VERY IMPORTANT]**

0 1 2 3 4

00 NOT AT ALL IMPORTANT
01
02
03
04 VERY IMPORTANT

Next I'd like to try to quantify the impact of the FlexTech Program. I'm going to ask you some questions about the following measures: **[LIST FR_A-FR_E]**.

[IF FR_[X]_Q=2 OR IF FR_[X]_Q=6, ASK FR6, ELSE SKIP TO FR7] (WHERE [X]=[A-E])

FR6. What is the likelihood that you would have incorporated **[FR_A-FR_E]** with the same high level of efficiency if your firm had not received the FlexTech Program's report or recommendations?

01 Definitely **would not** have incorporated measure of the same high level of efficiency (0%)
02 **May have** incorporated measure of the same high level of efficiency, even without the program. About what percent likelihood? _____% **[ACCEPT 0-100, EXCLUDING 0 AND 100]**
03 Definitely **would** have incorporated measure of the same high level of efficiency anyway (100%)
94 NOT YET INSTALLED
96 REFUSED
97 DON'T KNOW

[IF FR_[X]_Q =2 OR FR_[X]_Q=7, ASK FR7, ELSE SKIP TO FR8.] (WHERE [X]=[A-E])

FR7. What percentage of these high efficiency **[FR_A-FR_E]** would you have incorporated if you had not received the FlexTech Program's report or recommendations?
01 [RECORD PERCENT. [ACCEPT 0-100]]
94 NOT YET INSTALLED
96 REFUSED
97 DON'T KNOW

[IF NECESSARY, OR IF THE FLOW OF THE INTERVIEW DICTATES, YOU MAY DERIVE THIS VALUE BY ASKING THE FOLLOWING QUESTIONS]

[IF FR7=97, ASK BOTH FR7_1 AND FR7_2]

FR7_1) What is the share of measures that would have been incorporated (at any efficiency)?
01 [RECORD PERCENT. [ACCEPT 0-100]]
96 REFUSED
97 DON'T KNOW

FR7_2) What is the share of incorporated measures that would have been high efficiency?
01 [RECORD PERCENT. [ACCEPT 0-100]]
96 REFUSED
97 DON'T KNOW

[FOR ALL OF FR7 – INCLUDE A TEXT FIELD FOR INTERVIEWER TO RECORD ANY NOTES OR VERBATIM RESPONSES]

[FILL IN EITHER THE "LIKELIHOOD" VALUE OR THE "SHARE OF MEASURES" VALUE OR BOTH VALUES FOR EACH RELEVANT MEASURE CATEGORY.]			
MEASURE NAME	…WOULD HAVE BEEN INCORPORATED (AT HIGH EFFICIENCY) WITHOUT THE FLEXTECH (TA) PROGRAM		
	LIKELIHOOD FR6		SHARE FR7
[FR_A]	FR6_a	AND/OR	FR7_a
[FR_B]	FR6_b	AND/OR	FR7_b
[FR_C]	FR6_c	AND/OR	FR7_c
[FR_D]	FR6_d	AND/OR	FR7_d
[FR_E]	FR6_e	AND/OR	FR7_e

FR8. Most new equipment and design strategies have to meet current energy standards. But let's just focus on the fact that some of your new equipment strategies have even higher efficiencies than standard new equipment, and this new higher efficiency equipment provides extra energy savings….

Overall, across all measures, what percent of these extra energy savings would have been achieved anyway, even if the FlexTech Program did not exist? Please provide a lower and upper bound, and then your best estimate. Remember, I'm asking only about the extra savings from incorporating high efficiency equipment instead of standard efficiency equipment.

[READ IF NECESSARY: For example, 50% means that half of the extra savings from the high efficiency equipment would have been achieved anyway.]

FR8a. Lower bound

01 [RECORD PERCENT. ACCEPT [0-100]]
96 REFUSED
97 DON'T KNOW

FR8b. Upper bound

01 [RECORD PERCENT. ACCEPT [0-100]]
96 REFUSED
97 DON'T KNOW

FR8c. Best estimate

01 [RECORD PERCENT. ACCEPT [0-100]]
96 REFUSED
97 DON'T KNOW

[FR8a ≤ FR8c ≤ FR8b]

SECTION: ISO – INSIDE SPILLOVER

ISO1. Did your experience with the FlexTech Program influence your firm to incorporate additional energy efficiency measures at this site that had not been recommended by the FlexTech Program or any other NYSERDA programs (*i.e.*, measures that would not have happened without the influence of the program)?

01 YES
02 NO
96 REFUSED
97 DON'T KNOW

[IF ISO1=01 – YES, ASK ISO2. IF ISO1≠01, ASK OSO1.]

ISO2. Did these additional measures at this site save natural gas?

01 YES
02 NO
96 REFUSED
97 DON'T KNOW

[IF ISO2=01 – YES, ASK ISO3, ELSE ASKISO5.]

ISO3. For the project we have been discussing that was assisted by FlexTech, the program estimated your natural gas savings to be **[IMP_MMBtu]**. Would you estimate the natural gas savings from these extra measures to be less than, similar to, or more than this?

01 Less than the FlexTech project. How much less, in percentage terms?
1 [RECORD % [ACCEPT 0-100, EXCLUDING 100]]
2 [RECORD ACTUAL SAVINGS IN MMBtu [ACCEPT <[IMP_MMBtu]]
3 DON'T KNOW
02 About the same savings as the FlexTech Program project
03 More than the FlexTech Program project. How much more, in percentage terms?
1 [RECORD % [ACCEPT >100, EXCLUDING 100]]
2 [RECORD ACTUAL SAVINGS IN MMBtu [ACCEPT >[IMP_MMBtu]]
3 DON'T KNOW
96 REFUSED
97 DON'T KNOW

[01 – READ IF NECESSARY: If the additional measures saved ¼ the amount of gas saved through the FlexTech program, then the savings would be 25%]

[03 - READ IF NECESSARY: If the additional measures saved 1½ times the amount of gas saved through the FlexTech program, then the savings would be 150%.]

ISO4. To the best of your knowledge, what percentage of the natural gas savings from these extra measures are due to the FlexTech Program? [READ IF NECESSARY: That is, about what percentage of these measure savings were installed due to the influence of the FlexTech program?]

01 [RECORD PERCENTAGE. [ACCEPT ≤100]]
96 REFUSED
97 DON'T KNOW

ISO5. Did the additional measures at this site save electricity?

01 YES
02 NO
96 REFUSED

97 DON'T KNOW

[IF ISO5=01 – YES, ASK ISO6, ELSE SKIP TO OSO1]

ISO6. For the project we have been discussing that was assisted by FlexTech, the program estimated your electricity savings from the project to be **[IMP_kWh]**. Would you estimate the electricity savings from these extra measures to be less than, similar to, or more than this?

01 Less than the FlexTech Program project. How much less, in percentage terms?
1 [RECORD % [ACCEPT 0-100, EXCLUDING 100]]
2 [RECORD ACTUAL SAVINGS IN kWh [ACCEPT <[IMP_kWh]]
3 DON'T KNOW
02 About the same savings as the FlexTech Program project
03 More than the FlexTech Program project. How much more, in percentage terms?
1 [RECORD % [ACCEPT >100, EXCLUDING 100]]
2 [RECORD ACTUAL SAVINGS IN kWh [ACCEPT >[IMP_kWh]]
3 DON'T KNOW
96 REFUSED
97 DON'T KNOW

[01 – READ IF NECESSARY: If the additional measures saved ¼ the amount of electricity saved through the FlexTech program, then the savings would be 25%]

[03 - READ IF NECESSARY: If the additional measures saved 1½ times the amount of electricity saved through the FlexTech program, then the savings would be 150%.]

ISO7. To the best of your knowledge, what percentage of the electricity savings from these extra measures would you say are due to your experience with the FlexTech Program?

01 [RECORD PERCENTAGE] **[100% OR LESS]**
96 REFUSED
97 DON'T KNOW

SECTION: OSO – OUTSIDE SPILLOVER

OSO1. Did your company implement any additional energy efficiency measures at other facilities in New York State (excluding Long Island)?

01 YES
02 NO
96 REFUSED
97 DON'T KNOW

[IF OSO1=01, ASK OSO2, ELSE SKIP TO ST1]

OSO2. Did your experience with the FlexTech Program influence you to incorporate additional energy efficiency measures at other facilities in New York State (excluding Long Island) that did not receive support from the FlexTech Program or any other NYSERDA programs?

01 YES
02 NO
96 REFUSED
97 DON'T KNOW

[IF OSO2=01, ASK OSO2a, ELSE SKIP TO ST1]

OSO2a. About how many other facilities were influenced (that did not participate in NYSERDA programs)?

01 [RECORD NUMBER. [ACCEPT 1-50]]
96 REFUSED
97 DON'T KNOW

OSO3. Did these additional measures at these other facilities save natural gas?

01 YES
02 NO
96 REFUSED
97 DON'T KNOW

[IF OSO3=01, ASK OSO4, ELSE SKIP TO OSO6]

OSO4. On average, would you estimate the natural gas savings from these other non-program facilities to be less than, similar to, or more than the **[IMP_MMBtu]** natural gas savings from the energy efficiency measures incorporated through the FlexTech Program project?

01 Less than the FlexTech Program project. How much less, in percentage terms?
1 [RECORD % [ACCEPT 0-100, EXCLUDING 100]]
2 [RECORD ACTUAL SAVINGS IN MMBtu [ACCEPT <[IMP_MMBtu]]
3 DON'T KNOW
02 About the same savings as the FlexTech Program project
03 More than the FlexTech Program project. How much more, in percentage terms?
1 [RECORD % [ACCEPT >100, EXCLUDING 100]]
2 [RECORD ACTUAL SAVINGS IN MMBtu [ACCEPT >[IMP_MMBtu]]
3 DON'T KNOW
96 REFUSED
97 DON'T KNOW

[01 – READ IF NECESSARY: If the additional measures saved ¼ the amount of gas saved through the FlexTech program, then the savings would be 25%]

[03 - READ IF NECESSARY: If the same measures were implemented in a facility twice as big, then savings would be 200%. Remember, we are looking for savings "on average" (not in aggregate) across the many buildings that might be affected.**]**

OSO5. To the best of your knowledge, what percentage of the savings from natural gas measures at these non-program facilities would you say are due to your experience with the FlexTech Program?

01 [RECORD PERCENTAGE] **[100% OR LESS]**
96 REFUSED
97 DON'T KNOW

OSO6. Did these additional measures at these other facilities save electricity?

01 YES
02 NO
96 REFUSED
97 DON'T KNOW

[IF OSO6=01, ASK OSO7, ELSE SKIP TO ST1]

OSO7. On average, would you estimate the electricity savings from these other non-program facilities to be less than, similar to, or more than the **[IMP_kWh]** electricity savings from the energy efficiency measures incorporated through the FlexTech Program project?

01 Less than the FlexTech Program project. How much less, in percentage terms?
1 [RECORD % [ACCEPT 0-100, EXCLUDING 100]]
2 [RECORD ACTUAL SAVINGS IN kWh [ACCEPT <[IMP_kWh]]
3 DON'T KNOW

02 About the same savings as the FlexTech Program project

03 More than the FlexTech Program project. How much more, in percentage terms?
1 [RECORD % [ACCEPT >100, EXCLUDING 100]]
2 [RECORD ACTUAL SAVINGS IN kWh [ACCEPT >[IMP_ kWh]]
3 DON'T KNOW

96 REFUSED

97 DON'T KNOW

[01 – READ IF NECESSARY: If the additional measures saved ¼ the amount of gas saved through the FlexTech program, then the savings would be 25%]

[03 - READ IF NECESSARY: If the same measures were implemented in a facility twice as big, then savings would be 200%. Remember, we are looking for savings "on average" (not in aggregate) across the many buildings that might be affected.**]**

OSO8. To the best of your knowledge, what percentage of the savings from the electricity measures at these non-program facilities would you say are due to your experience with the FlexTech Program?

01 [RECORD PERCENTAGE. [ACCEPT 0-100]]

96 REFUSED

97 DON'T KNOW

SECTION: ST – FIRMOGRAPHICS (STATISTICS)

[CATI INSTRUCTIONS: THIS SECTION DOES NOT NEED TO BE REPEATED FOR ANY ADDITIONAL DECISION MAKER SURVEYS FOR THE SAME COMPANY/ORGANIZATION]

ST1. What is the Principal Activity of the building in which the equipment was installed? [READ IF NECESSARY. MARK ALL THAT APPLY.]

01 EDUCATION
02 FOOD SALES
03 FOOD SERVICE
04 HEALTH CARE
05 LODGING
06 RETAIL/MERCANTILE
07 OFFICE
08 PUBLIC ASSEMBLY
09 PUBLIC ORDER AND SAFETY
10 RELIGIOUS WORSHIP
11 SERVICE
12 WAREHOUSE AND STORAGE
13 MANUFACTURING (In what industry? – E.G. CHEMICAL, FOOD, PAPER, ETC.)
14 VACANT
95 OTHER (SPECIFY)

96 REFUSED
97 DON'T KNOW

ST2. Approximately, when was this building built? [READ IF NECESSARY.]
01 Before 1960
02 1961-1970
03 1971-1980
04 1981-1990
05 1991-2000
06 2001-2005
07 After 2005
96 REFUSED
97 DON'T KNOW

ST3. What is the approximate square footage of the building where the equipment was installed? [READ IF NECESSARY: If the equipment was installed in multiple buildings, please estimate the total square footage across all buildings.]
01 Less than 1,000 square feet
02 1,000 to 4,999
03 5,000 to 14,999
04 15,000 to 24,999
05 25,000 to 49,999
06 50,000 to 99,999
07 100,000 to 199,999
08 200,000 to 499,999
09 500,000 or more
96 REFUSED
97 DON'T KNOW

ST4. How many employees does your firm have?
01 Fewer than 5
02 5 to 9
03 10 to 19
04 20 to 49
05 50 to 99
06 100 to 249
07 250 or More
96 REFUSED
97 DON'T KNOW

ST5. Is your company independent, or part of a larger company?
01 Independent
02 Part of a larger company
03 OTHER (Specify)
96 REFUSED
97 DON'T KNOW

ST6. How many locations/establishments does your firm have?
01 One
02 2 to 5
03 6 to 10
04 11 to 20
05 More than 20

96 REFUSED
97 DON'T KNOW

That completes this telephone survey.

[IF MAR_ADD==1, READ]

We also wanted to let you know that Energy Resource Solutions (ERS) will be calling again in the next few months to conduct a follow-up Measure Adoption Rate survey. This additional survey should take no more than 10 minutes.

THANK YOU FOR YOUR TIME!

Appendix E
TA Service Provider Net-to-Gross Telephone Questionnaire

NYSERDA Flexible Technical Assistance (FlexTech)
Program Impact Evaluation 2011 Service Provider Survey
Final for CATI Programming
4/27/2011

Hello may I please speak to **[NAME1]**? I'm calling on behalf of the New York State Energy Research and Development Authority (or NYSERDA). We are conducting research for NYSERDA's FlexibleTechnical Assistance Program (also known as FlexTech).

[IF INTRO==01, READ INTRO1, ELSE SKIP TO INTRO2][ONE ACCOUNT SITE]
INTRO 1

Our records have [**COMPANY1**] listed as the Technical Assistance Service Provider for [**ACCOUNT_SITE**] and indicate that you may be the best contact at your firm for this project. I'd like to ask you some questions regarding the decision-making process for this project or projects similar to this one.

[READ IF NECESSARY: We are researching a small carefully designed sample of projects that participated in the NYSERDA FlexTech Program. Because we are only talking to a few people, your participation in this evaluation is very important to us. The information you provide will be used to assess program accomplishments and improve NYSERDA's programs. Your responses will be kept confidential to the extent permitted by law.]

[GO TO Q1]

[IF INTRO==02, READ INTRO2][MULTIPLE ACCOUNT SITES]
INTRO 2

We would like to conduct interviews regarding the decision-making process for the following projects:

[**HAVE MULTIPLE SITE LISTINGS ON PAPER SAMPLE]**Our records have [**COMPANY1**] listed as the Technical Assistance Service Provider for these projects and indicate that you may be the best contact at your firm regarding these projects. I'd like to ask you some questions regarding the decision-making process for these projects or projects similar to these.

[READ IF NECESSARY: We are researching a small carefully designed sample of projects that participated in the NYSERDA FlexTech Program. Because we are only talking to a few people, your participation in this evaluation is very important to us. The information you provide will be used to assess program accomplishments and improve NYSERDA's programs. Your responses will be kept confidential to the extent permitted by law.]

[GO TO Q1]

Q1. Are you the appropriate person to discuss issues related to the FlexTech project at [**ACCOUNT SITE**]? **[HAVE MULTIPLE SITE LISTINGS ON PAPER SAMPLE.]**

01 YES
02 NO, NOT CORRECT RESPONDENT
98 REFUSED
99 DON'T KNOW

[IF Q1=01, SKIP TO SCR-1, ELSE ASK Q2]

Q2. Can you provide me with a contact name and phone number for a person who can speak about the project at [**ACCOUNT SITE**]? **[REPEAT FOR ALL ACCOUNT SITES]**

01 YES [**SPECIFY NAME, NUMBER**]
02 NO **[TERMINATE]**
96 REFUSED **[TERMINATE]**
97 DON'T KNOW **[TERMINATE]**

[LIST NAME AND PHONE # IN SPACE BELOW]

Name: _____

Phone: (____)_____ext._____

[CONTACT THIS PERSON, REPEAT INITIAL INTRODUCTION AND CONTINUE WITH THE FOLLOWING QUESTION]

[SCHEDULING]

SCR-1. This survey will take about 20 minutes to complete. Can we discuss the project now, or can we schedule a time when I can call you back?

1. CAN DISCUSS NOW **[PROCEED TO INFLUENCE OF AUDIT]**
2. CALL BACK ON:_____ AT TIME:_____

[CONDUCT AN INTERVIEW FOR EACH ACCOUNT SITE]

[CONDUCT SURVEY AUD1-FR8c AND OSO1 FOR EACH ACCOUNT SITE]

INFLUENCE OF AUDIT

A couple of months ago we conducted a Measure Adoption Rate interview regarding equipment adoption relating to the project at **[ACCOUNT_SITE]** that participated in NYSERDA's Flexible Technical Assistance, also known as FlexTech**.** That prior survey provided us with current information as to what recommendations were adopted and differences between what was adopted versus what was recommended.

AUD1. If the FlexTech Program or its assistance had not been available, what is the likelihood that this customer would have performed a study of the same or similar quality than the study performed under the FlexTech Program for [**ACCOUNT_SITE**]? Please rate the likelihood in percentage terms with 0% being not likely at all and 100% being extremely likely.

01 **[[RECORD PERCENT. ACCEPT [0-100]]**
96 REFUSED
97 DON'T KNOW

AUD2. How comprehensive was the study or audit you conducted for this project through FlexTech compared to those you conduct outside the FlexTech program? Would you say this project's study or audit conducted through FlexTech was…?

01 More comprehensive or extensive
02 Similar in comprehensiveness
03 Less comprehensive or extensive
96 REFUSED
97 DON'T KNOW

SECTION: FR – FREE RIDERSHIP

The equipment adoption interview conducted last fall found that **[ACCOUNT SITE]** made at least partial adoption of **the following measures**: **[LIST MEASURES FR_A-FR_E]**

FR1. Prior to participating in the FlexTech Program, were there plans to install any of the adopted energy efficiency or demand measures at your customer's site that were recommended by the FlexTech report?

01 YES
02 NO
96 REFUSED
97 DON'T KNOW

[IF FR1=01, ASK FR2, ELSE SKIP TO FR3]

FR2. Could you please describe any plans that your customer had to incorporate the adopted measures prior to participating in the FlexTech Program?

[RECORD VERBATIM]

[BASED ON RESPONSES TO FR2, FILL IN A "0 TO 4" SCORE INDICATING THE EXTENT TO WHICH RESPONDENT WAS ALREADY PLANNING TO INCORPORATE THE ENERGY EFFICIENCY MEASURE. DO NOT ASK RESPONDENT DIRECTLY. "0" INDICATES THAT RESPONDENT HAD NO PLANS AT ALL; "4" INDICATES THAT RESPONDENT HAD DOCUMENTED PLANS AND HAD BUDGETED FOR ALL OF THE EFFICIENCY MEASURES.]

[NO PLANS] **[DOCUMENTED PLANS/BUDGET]**

0 1 2 3 4

[FULL GUIDELINES FOR ASSIGNING HIGH-EFFICIENCY PROJECT PLANNING SCORE]

SCORE	EXTENT OF PLANNING
0	NO PLANS FOR HIGH EFFICIENCY EQUIPMENT; RESPONDENT MAY HAVE CONSIDERED ALTERNATIVE TECHNOLOGY OPTIONS, BUT DID NOT EXPLICITLY CONSIDER HIGH EFFICIENCY.
1	INITIAL STEPS TOWARD CONSIDERATION OF HIGH EFFICIENCY SUCH AS REQUESTING INFORMATION ON OR DISCUSSING, IN GENERAL, HIGH EFFICIENCY OPTIONS WITH VENDORS OR CONTRACTORS.
2	IN-DEPTH DISCUSSION OR CONSIDERATION OF SPECIFIC TYPES OF HIGH EFFICIENCY EQUIPMENT (E.G., LIGHTING, HVAC, APPLIANCES), INCLUDING THEIR POSITIVE AND NEGATIVE ATTRIBUTES AND COSTS.
3	IDENTIFICATION OF SPECIFIC EQUIPMENT MANUFACTURERS AND MODELS, INCLUDING ASSESSMENT OF THEIR RELATIVE COSTS AND PERFORMANCE CHARACTERISTICS.
4	HIGH EFFICIENCY EQUIPMENT AND DESIGNS FULLY SPECIFIED AND EXPLICITLY SELECTED OR INCORPORATED INTO PROJECT BUDGET.

FR3. Did the project's participation in the FlexTech Program in any way influence the type, efficiency level, or amount of measures that were incorporated?

01 YES
02 NO (all equipment would have been incorporated at the same high efficiencies)
96 REFUSED
97 DON'T KNOW

[IF FR3=01 – YES, ASK FR4, ELSE SKIP TO FR5]

FR4. Please briefly describe how you think the FlexTech Program influenced the decision to incorporate high efficiency measures at [**ACCOUNT SITE**].

[BASED ON RESPONSE TO FR4, FILL IN A "0 TO 4"SCORE INDICATING THE EXTENT TO WHICH THE PROGRAM INFLUENCED THE DECISION TO INCORPORATE HIGH EFFICIENCY MEASURES. DO NOT ASK RESPONDENT DIRECTLY. "0" INDICATES THAT THE PROGRAM HAD NO INFLUENCE; "4" INDICATES THAT THE PROGRAM WAS THE PRIMARY REASON THAT HIGH EFFICIENCY MEASURES WERE INCORPORATED.]

[NO PROGRAM INFLUENCE] **[PROGRAM PRIMARY INFLUENCE]**

0 1 2 3 4

[FULL GUIDELINES FOR ASSIGNING PROGRAM INFLUENCE SCORE]

SCORE	CHARACTERIZATION OF PROGRAM INFLUENCE
0	NO INFLUENCE ON THE DECISION TO INSTALL HIGH EFFICIENCY EQUIPMENT. ALL EQUIPMENT WOULD HAVE BEEN INSTALLED AT THE SAME EFFICIENCIES EVEN WITHOUT THE PROGRAM.
1	PROGRAM HELPED IN MAKING FINAL DECISION ON EQUIPMENT THAT HAD ALREADY BEEN THOROUGHLY CONSIDERED.
2	PROGRAM LENT CREDIBILITY TO THE DECISION TO INVEST IN HIGH EFFICIENCY AND/OR IT PROVIDED INFORMATION THAT HELPED EXPAND THE QUANTITY, SCOPE, OR EFFICIENCY OF THE EQUIPMENT.
3	PROGRAM IDENTIFIED A SIGNIFICANT NUMBER OF SPECIFIC HIGH EFFICIENCY OPTIONS THAT WERE INSTALLED BUT THAT HAD NOT PREVIOUSLY BEEN CONSIDERED AND/OR PROGRAM WAS A MAJOR DRIVER BEHIND A SIGNIFICANT INCREASE IN THE QUANTITY, SCOPE, OR EFFICIENCY OF HIGH EFFICIENCY EQUIPMENT.
4	PROGRAM WAS THE PRIMARY REASON THAT HIGH EFFICIENCY EQUIPMENT WAS INSTALLED IN THE PROJECT.

FR5. On a scale of 0 to 4, where 0 = "not at all important" and 4 = "very important," please indicate how important you think the FlexTech Program was in [**ACCOUNT SITE**] decision to incorporate high efficiency measures at this site?

[NOT AT ALL IMPORTANT] **[VERY IMPORTANT]**

0 1 2 3 4

05 NOT AT ALL IMPORTANT
06

07
08
09 VERY IMPORTANT

Next I'd like to try to quantify the impact of the FlexTech Program at [**ACCOUNT SITE**]. Let me ask about the **measures I listed previously**.

[IF FR_[X]_Q=2 OR IF FR_[X]_Q=6, ASK FR6, ELSE SKIP TO FR7] (WHERE [X]=[A-E])

FR6. What is the likelihood that **[FR_A-FR_E]** of the same high level of efficiency would have been incorporated at this site if it had not been for the FlexTech Program and its assistance?

01 Definitely **would not** have incorporated measure of the same high level of efficiency (0%)
02 **May have** incorporated measure of the same high level of efficiency, even without the program. About what percent likelihood? _____% **[ACCEPT 0-100, EXCLUDING 0 AND 100]**
03 Definitely **would** have incorporated measure of the same high level of efficiency anyway (100%)
94 NOT YET INSTALLED
96 REFUSED
97 DON'T KNOW

[IF FR_[X]_Q =2 OR FR_[X]_Q=7, ASK FR7, ELSE SKIP TO FR8.] (WHERE [X]=[A-E])

FR7. What percentage of these high efficiency **[FR_A-FR_E]** would you have incorporated if you had not received the FlexTech Program's report or recommendations?

02 [RECORD PERCENT. [ACCEPT 0-100]]
94 NOT YET INSTALLED
98 REFUSED
99 DON'T KNOW

[READ IF NECESSARY: So, assuming that you had decided to incorporated [FR_A-FR_E], what share or percent of the measures would you have implemented? That is, would you have made all of the changes or only part of them? And if part, what percent would you say you would have implemented?]

[IF FR7=97, ASK FR7_1 AND FR7_2]

FR7_1) What is the share of measures that would have been incorporated (at any efficiency)?

02 [RECORD PERCENT. [ACCEPT 0-100]]
98 REFUSED
99 DON'T KNOW

FR7_2) What is the share of incorporated measures that would have been high efficiency?

02 [RECORD PERCENT. [ACCEPT 0-100]]
98 REFUSED
99 DON'T KNOW

[FILL IN EITHER THE "LIKELIHOOD" VALUE OR THE "SHARE OF MEASURES" VALUE OR BOTH VALUES FOR EACH RELEVANT MEASURE CATEGORY.]

MEASURE NAME	…WOULD HAVE BEEN INCORPORATED (AT HIGH EFFICIENCY) WITHOUT THE FLEXTECH (TA) PROGRAM		
	LIKELIHOOD FR6		SHARE FR7
[FR_A]	FR6_a	AND/OR	FR7_a
[FR_B]	FR6_b	AND/OR	FR7_b
[FR_C]	FR6_c	AND/OR	FR7_c
[FR_D]	FR6_d	AND/OR	FR7_d
[FR_E]	FR6_e	AND/OR	FR7_e

FR8. Most new equipment and design strategies have to meet current energy standards. But let's just focus on the fact that some of the new equipment, incorporated as a result of the FlexTech studies have even higher efficiencies than standard new equipment, and this new higher efficiency equipment provides extra energy savings.

Overall, across all measures, what percent of these extra energy savings at [ACCOUNT SITE] would have been achieved anyway, even if the FlexTech Program did not exist? Please provide a lower and upper bound, and then your best estimate.

[READ IF NECESSARY: For example, 50% means that half of the extra savings from the high efficiency equipment would have been achieved anyway.]

FR8a. Lower bound

02 [RECORD PERCENT. ACCEPT [0-100]]
98 REFUSED
99 DON'T KNOW

FR8b. Upper bound

02 [RECORD PERCENT. ACCEPT [0-100]]
98 REFUSED
99 DON'T KNOW

FR8c. Best estimate

02 [RECORD PERCENT. ACCEPT [0-100]]
98 REFUSED
99 DON'T KNOW

[FR8a ≤ FR8c ≤ FR8b]

SECTION: NPS1 – INDICATOR OF NON-PARTICIPANT SPILLOVER

NPS1. Do you think that other engineering firms that are **not** participating in the FlexTech Program have **increased**, **decreased** or **not changed** the extent to which they incorporate energy efficient measures?

01 INCREASED
02 DECREASED
03 NOT CHANGED
96 REFUSED
97 DON'T KNOW

[IF NPS2=01 OR 02, ASK NPS2, ELSE SKIP TO OSO1]

NPS2. Have they [increased/decreased] their incorporation of energy efficient measures a lot or a little?

01 INCREASED A LOT
02 INCREASED A LITTLE
03 DECREASED A LITTLE
04 DECREASED A LOT
96 REFUSED
97 DON'T KNOW

[IF NPS2=01 OR 02, ASK NPS3, ELSE SKIP TO OSO1]

NPS3. Was the FlexTech Program a **major**, **minor**, or **not a factor** in this increased incorporation of energy efficiency measures? [READ IF NECESSARY: That is, was the Flextech program a major factor, minor factor or not a factor in increasing the use of energy efficiency measures by these other firms who are not participating in the FlexTech program.]

01 MAJOR
02 MINOR
03 NOT A FACTOR
96 REFUSED
97 DON'T KNOW

SECTION: OSO – OUTSIDE SPILLOVER

OSO1. To your knowledge, did your customer on this project implement any additional energy efficiency measures at other facilities in New York State (excluding Long Island) that did not receive assistance from FlexTech or any other NYSERDA Program?

01 YES
02 NO
96 REFUSED
97 DON'T KNOW

Now, I'd like to ask some questions about your firm's experience and interaction with energy efficiency as a result of your participation in NYSERDA's FlexTech Program.

OSO2. Did your experience with the FlexTech Program influence **your firm** to incorporate additional energy efficiency measures or designs at other facilities in New York State (excluding Long Island) that did not receive support from the FlexTech Program or any other NYSERDA programs?

03 YES
04 NO

98 REFUSED
99 DON'T KNOW

[IF OSO2=01, ASK OSO2a, ELSE SKIP TO P1]

OSO2a. About how many other facilities were influenced (that did not participate in NYSERDA programs)?

02 [RECORD NUMBER. [ACCEPT 1-95]]
98 REFUSED
99 DON'T KNOW

OSO3. Did these additional measures at these other facilities save natural gas?

03 YES
04 NO
98 REFUSED
99 DON'T KNOW

[IF OSO3=01, ASK OSO4, ELSE SKIP TO OSO6]

OSO4. On average, would you estimate the natural gas savings from these other non-program facilities to be less than, similar to, or more than the **[AVG_MMBtu]** natural gas savings from the energy efficiency measures incorporated through the FlexTech Program project?

[INTERVIEWER NOTE: For engineering firms with more than one FlexTech project, this number is the average savings across **all** their FlexTech projects.]

04 Less than the FlexTech Program project. How much less, in percentage terms?
   a. [RECORD % [ACCEPT 0-100, EXCLUDING 100]]
   b. [RECORD ACTUAL SAVINGS IN MMBtu [ACCEPT <[IMP_MMBtu]]
   97 DON'T KNOW
05 About the same savings as the FlexTech Program project
06 More than the FlexTech Program project. How much more, in percentage terms?
   a. [RECORD % [ACCEPT >100, EXCLUDING 100]]
   b. [RECORD ACTUAL SAVINGS IN MMBtu [ACCEPT >[IMP_MMBtu]]
   97 DON'T KNOW
98 REFUSED
99 DON'T KNOW

[01 – READ IF NECESSARY: If the additional measures saved ¼ the amount of gas saved through the FlexTech Program, then the savings would be 25%.]

[03 – READ IF NECESSARY: If the same measures were implemented in a facility twice as big, then savings would be 200%. Remember, we are looking for savings "on average" (not in aggregate) across the many buildings that might be affected.**]**

OSO5. To the best of your knowledge, what percentage of the savings from natural gas measures at these non-program facilities would you say are due to your experience with the FlexTech Program?

02 [RECORD PERCENTAGE [ACCEPT 0-100]]
98 REFUSED
99 DON'T KNOW

OSO6. Did these additional measures at these other facilities save electricity?

03 YES
04 NO
98 REFUSED

99 DON'T KNOW

[IF OSO6=01, ASK OSO7, ELSE SKIP TO P1]

OSO7. On average, would you estimate the electricity savings from these other non-program facilities to be less than, similar to, or more than the _____ **[AVG_kWh]** electricity savings from the energy efficiency measures incorporated through the FlexTech Program project?

04 Less than the FlexTech Program project. How much less, in percentage terms?
  01 [RECORD % [ACCEPT 0-100, EXCLUDING 100]]
  02 [RECORD ACTUAL SAVINGS IN kWh [ACCEPT <[IMP_kWh]]
  97 DON'T KNOW
05 About the same savings as the FlexTech Program project
06 More than the FlexTech Program project. How much more, in percentage terms?
  01 [RECORD % [ACCEPT >100, EXCLUDING 100]]
  02 [RECORD ACTUAL SAVINGS IN MMBtu [ACCEPT >[IMP_MMBtu]]
  97 DON'T KNOW
98 REFUSED
99 DON'T KNOW

[01 – READ IF NECESSARY: If the additional measures saved ¼ the amount of gas saved through the FlexTech Program, then the savings would be 25%.]

[03 - READ IF NECESSARY: If the same measures were implemented in a facility twice as big, then savings would be 200%. Remember, we are looking for savings "on average" (not in aggregate) across the many buildings that might be affected.**]**

OSO8. To the best of your knowledge, what percentage of the savings from the electricity measures at these non-program facilities would you say are due to your experience with the FlexTech Program?

02 [RECORD PERCENTAGE. [ACCEPT 0-100]]
98 REFUSED
99 DON'T KNOW

OTHER PROGRAMS

P1. To the best of your knowledge, has your organization participated in any other NYSERDA or **New York Energy $mart**[SM] programs in the past three years either yourself or on behalf of a client?

01 Did not participate in any other NYSERDA program(s);
02 Participated in other NYSERDA program(s) but can't recall which one(s);
03 Participated in NYSERDA program(s):
96 REFUSED
97 DON'T KNOW

[IF P1=03, ASK P2, ELSE SKIP TO ST1]

P2. Which NYSERDA programs did you participate in? [MARK ALL THAT APPLY. PROBE TO CODE.]

01 Commercial / Industrial Performance Program (CIPP)
02 Peak-Load Reduction Program (PLRP)
03 Smart Equipment Choices Program (SEC)

04 New York Energy $mart[sm] Loan Fund (Loan Fund)
05 Small Commercial Lighting Program (SCLP)
06 End-Use Renewables Program
95 Other Program (SPECIFY:_____)
96 REFUSED
97 DON'T KNOW

SECTION: ST – FIRMOGRAPHICS (STATISTICS)

ST1. Is your firm a(n) . . . ?
01 ESCO
02 Architectural firm
03 Engineering firm
04 HVAC contractor
95 Or something else? (Specify: _____)
96 REFUSED
97 DON'T KNOW

ST2. What percentage of all your projects address the following areas? [READ IF NECESSARY: If half of all the projects your firm works on addresses lighting, then the response for lighting should be 50%. If all of your projects address lighting, then the response should be 100%]
[NOTE: DOES NOT NEED TO ADD UP TO 100%]
_____ % Lighting
_____ % HVAC
_____ % Motors and drives
_____ % Building Shell
_____ % Load management/curtailment
_____ % CHP (Combined Heat and Power)
_____ % Process improvements (manufacturing, and/or water and wastewater)
_____ % Other _____
96 REFUSED
97 DON'T KNOW

ST3. What percentage of your project work is spent performing the following activities (based on the number of projects that you have worked on)? [**READ ENTIRE LIST THEN RECORD – VERIFY THAT TOTAL EQUALS 100%**]
_____ % Studies/audits
_____ % Design
_____ % Install
_____ % Maintenance/Commissioning
_____ % Other (specify: _____)
96 REFUSED
97 DON'T KNOW

ST4. We would like to know the four facility types in which your firm works most frequently in New York State (not including Long Island). I have a list of facility types I will go through, and then I will work with you, so you can identify the top four.

[WITHIN TOP FOUR]
a. _____

b. _____
c. _____
d. _____

01 Agriculture
02 Education, Schools, colleges, libraries, laboratories (non-manufacturing owned)
03 Grocery / food sales
04 Food Service, restaurants
05 Government Service Building
06 Health Care, hospitals and other health treatment
07 Lodging, Hotel, Motel
08 Retail / Mercantile
09 Multifamily
10 Office and bank building
11 Warehouse and Storage (excluding manufacturer owned)
12 Manufacturing plants, warehouse, laboratories (Identify Industry Type
_____ (e.g., chemical, food, paper, etc.)
13 Water / wastewater
14 Public buildings, churches, amusement / social or other assembly buildings
15 Other miscellaneous non-residential buildings
95 NONE / NO OTHER
96 REFUSED
97 DON'T KNOW

We would like to know what share of your work are in each of these top four facility types? What percent of your firm's work in New York are projects for [INSERT ST4a-ST4d]?

e. [PERCENT CORRESPONDING TO ST4a]
f. [PERCENT CORRESPONDING TO ST4b]
g. [PERCENT CORRESPONDING TO ST4c]
h. [PERCENT CORRESPONDING TO ST4d]

996 REFUSED
997 DON'T KNOW

ST6. Approximately how many full time employees or full time equivalents does your organization have at all your locations in New York State? [READ IF NECESSARY: In New York State, excluding Long Island]

01 Fewer than 5
02 5 to 9
03 10 to 19
04 20 to 49
05 50 to 99
06 100 to 249
07 250 or More
96 REFUSED
97 DON'T KNOW

ST7. Is your company independent, or part of a larger company?

01 Independent
02 Part of a larger company
03 OTHER (Specify: _____)
96 REFUSED

97 DON'T KNOW

ST8. How many locations/establishments does your firm have? [READ IF NECESSARY: In New York State, excluding Long Island]

01 One
02 2 to 5
03 6 to 10
04 11 to 20
05 More than 20
96 REFUSED
97 DON'T KNOW

ST9. Do you have any other comments about NYSERDA's FlexTech Program?

01 [RECORD VERBATIM]
96 REFUSED
97 DON'T KNOW

That completes this telephone survey. Thank you very much for your assistance!

Appendix F
Tracking System Description

The FlexTech evaluation team used the web-based tool Salesforce to track the telephone survey calls, document the survey results, and track the project's progress. This appendix describes that tool in more detail.

TRACKING SYSTEM – PROGRAM DATA ENTRY

After the selection of the MAR sample, the NYSERDA supplied data set was sorted using Microsoft Access and Excel. The tracking system required that all projects have a unique identifier so that data can be updated and reported correctly. Therefore, each project was given a name, which was a combination of the evaluation name (NYTA), data set version, and NYSERDA project number (PO ID). Similarly, project measures were given a unique identifier, which was a combination of the NYSERDA project number and a unique four-digit number.

Prior to uploading the data into the tracking system, quality assurance checks were completed by evaluators to insure that data was not lost or incorrectly mapped to the wrong project. A random selection of the import data set was compared to the raw data set provided by NYSEDA. A similar quality check was performed after the data was uploaded. Project sampling information, measure data, and contact information was checked for consistency.

TRACKING SYSTEM – KEY FIELDS

The data imported into the tracking system was mapped to the fields that were described in the MAR survey instrument. In order to maintain consistency with the NYSERDA database and sampling strategy, some fields contained data that could not be edited (e.g., the FlexTech study's savings projections, the measure category, and the upstate/downstate classification). These fields were locked so that users could not accidentally change the values after the data import. In contrast, most fields were editable so that users could insert survey responses and track their progress.

Figure 1 illustrates the main account page to which all contacts and efficiency measures are linked. On this page, the high level status of the project was tracked using the "Account Status" and "Stage" fields. These fields contain dropdown options that were used to track the project's progress from "Sample – None" stage to the "Telephone Survey – Complete" or "Site Evaluation – Complete" stage. This page also tracked the "Site Dropped?" field. These fields were used to record when and why a site was dropped.

Figure 1. Tracking System Main Account Page

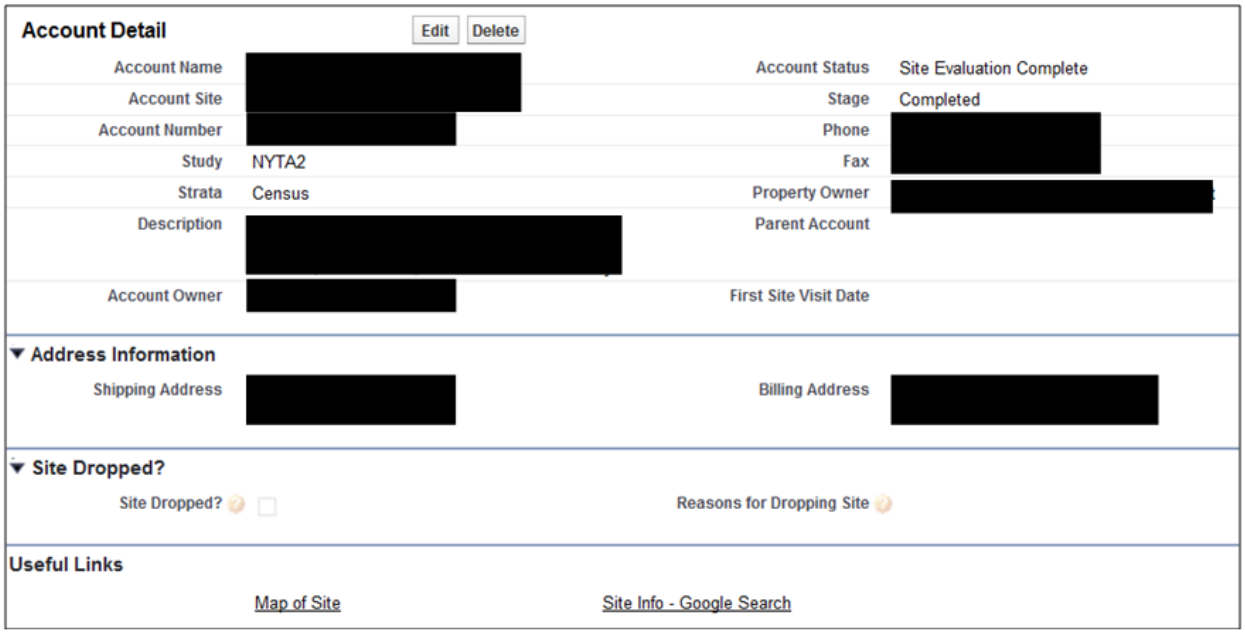


Figure 2 illustrates the FlexTech account information page, which includes the sampling details, MAR survey questions related to project funding, and site visit details if applicable (left blank if the project was not in the on-site visit sample).

Figure 2. FlexTech Account Information Page

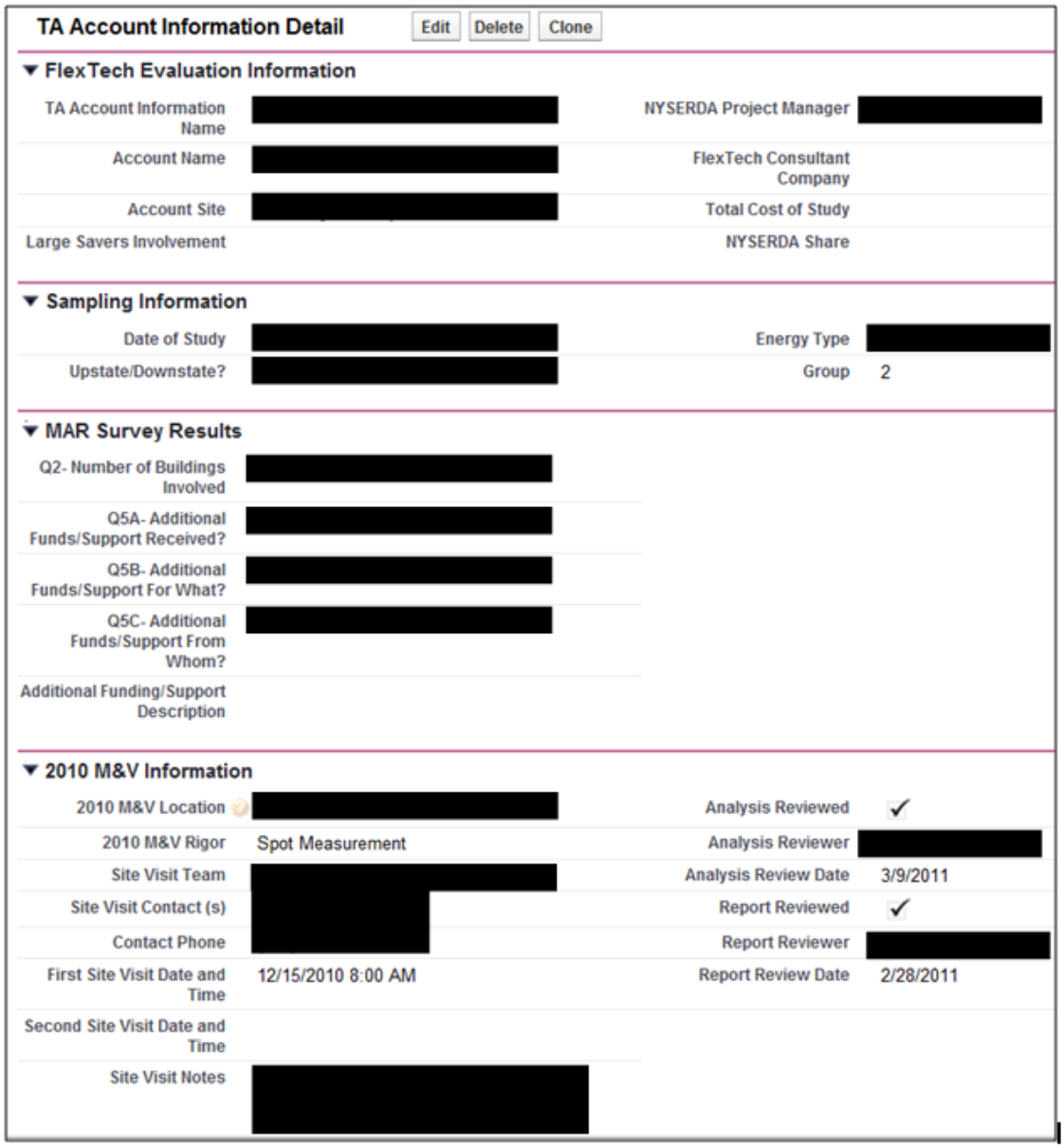


Figure 3 illustrates the contact detail page. Of particular importance are the "Decision-Maker" checkbox and the "Contact Result" field. The decision-maker field was used to identify who should be contacted for the net-to-gross surveys, while the contact result was used to determine which contacts completed the MAR surveys. Also note the "Do Not Call" checkbook on the right side of the page, which was used to tag contacts that either indicated a preference to not be called in the future or were deemed to be incorrect for the MAR calls. Prior to placing any calls, surveyors verified that the "Do Not Call" checkbox was empty.

Figure 3. Contact Detail Page

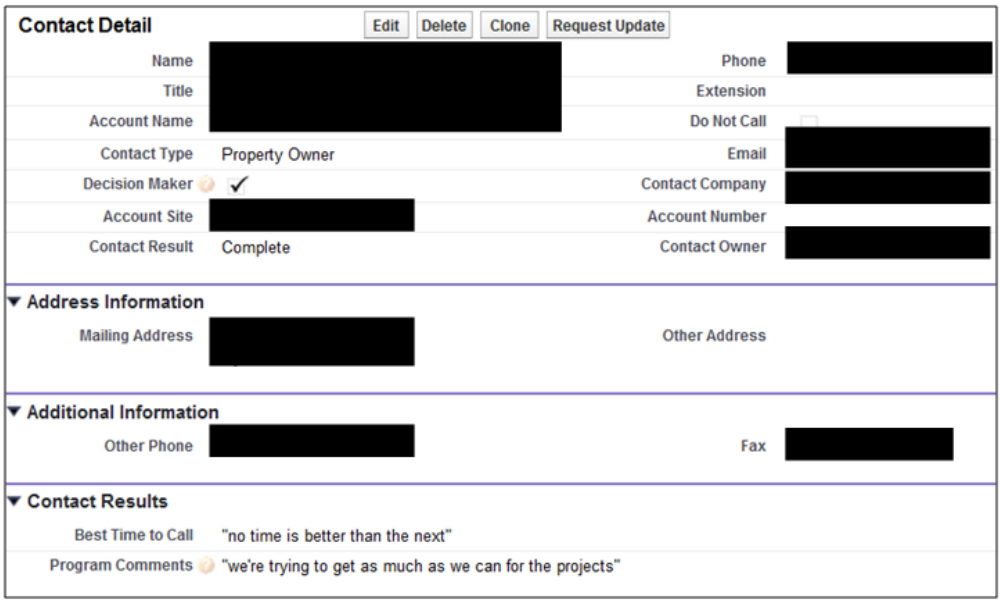
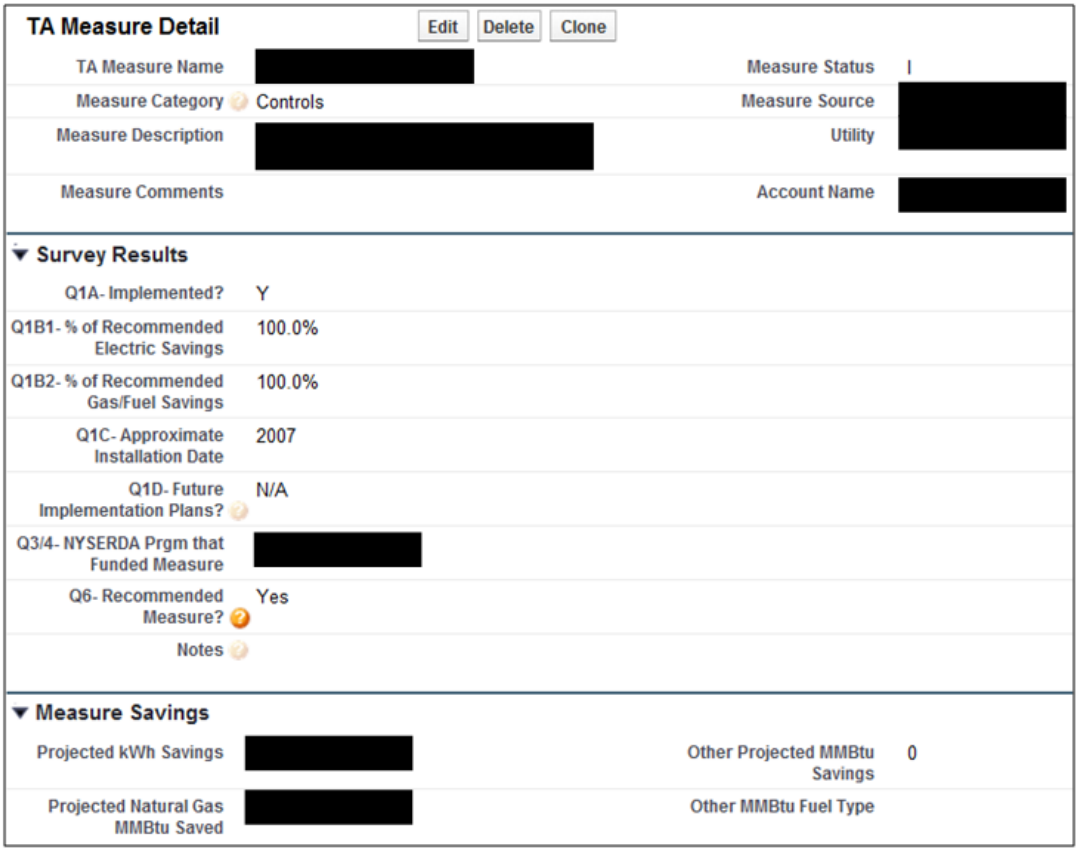


Figure 4 illustrates the measure details. Each measure has its own page that includes information from the FlexTech study and the measure-related survey responses. The savings projections were provided by NYSERDA, while the fields listed under the "Survey Results" section header were obtained during the MAR surveys.

Figure 4. Measure Detail Page



Prior study responses were imported into the tracking system as separate measures within each project, identified by a "Prior Study Noncogen" or "Prior Study Nonelectric" measure source field. Prior to placing any calls, surveyors reviewed all measure details to determine if the prior responses could be used in lieu of survey responses.

How the Tracking System Tracked Progress

To make the survey preparation process easier, the tracking system was used to generate MAR survey documents that were pre-populated with facility information (account number, address, NYSERDA project manager, etc.). This saved time by allowing the surveyors to bypass the tedious process of customizing each survey instrument. Additionally, the tracking system generated an Excel spreadsheet that listed all measures for a particular project. The table included the following measure details that were provided by NYSERDA:

- Measure category (e.g., Controls, HVAC, Lighting, etc.)
- Measure description
- Measure status (I – implemented, R – recommended, RS – recommended for further study, NR – not recommended, RME – recommended mutually exclusive, ME – mutually exclusive, RNE – recommended non-energy)
- Projected kWh savings
- Projected MMBtu savings

The table also included spaces for the following MAR survey responses (spaces were filled in for measures answered in the prior study):

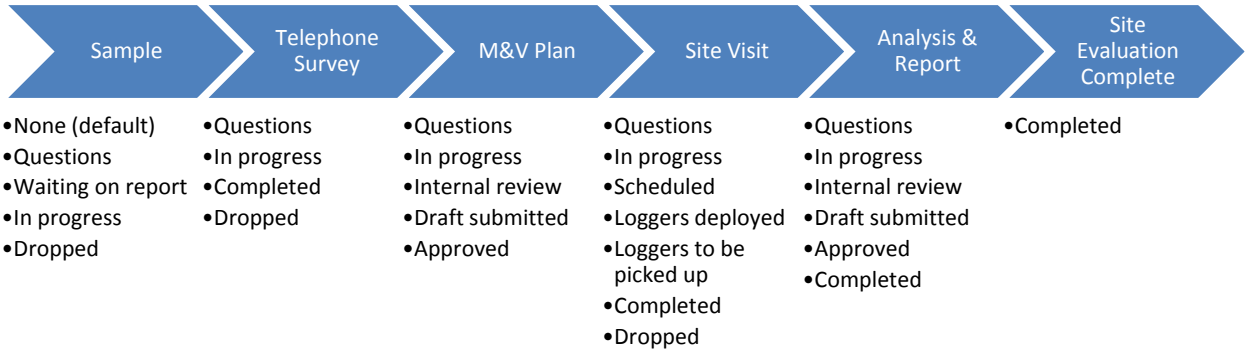
- Was the measure implemented? (Yes, No, Partial, In Progress)
- What percent of the recommended electric savings is achieved? (kWh MAR)
- What percent of the recommended natural gas savings is achieved? (MMBtu MAR)
- When was the measure installed?
- If the measure was not installed completely, do you plan to install it in the future?
- What NYSERDA program funded the measure?

These tables were used by surveyors to quickly view and record survey responses to be imported into the tracking system. The written documents also provided space for notes, which proved very helpful to refresh evaluators' memories when reviewing projects months after MAR surveys were completed.

Each contact attempt was logged in the tracking system. Evaluators entered notes about call results and emails, which could be accessed later. The tracking system also enabled surveyors to make tasks for future contact attempts.

Progress was tracked for each project using the "Account Status" and "Stage" fields on the account page. Figure 5 shows the progression of a project from the import phase to the site evaluation completion.

Figure 5. Project Progression from the Import Phase to the Site Evaluation Completion



A date and time stamp was recorded each time one of these tracking fields was changed. This allowed project managers to see the progression of the survey and cite evaluation process as a whole.

Tracking System Reports

Weekly meetings included progress updates and charts that were generated with the tracking system. All project statuses could be reported on in real-time as illustrated in figures 6 and 7. Figure 6 indicates the number of projects in the Telephone Survey phase versus a Sample. The Sample phase indicated projects that were still under review. Figure 7 indicates the status of projects assigned to our engineers at a given moment in time.

Figure 6. Number of Projects in the Telephone Survey Phase versus a Sample

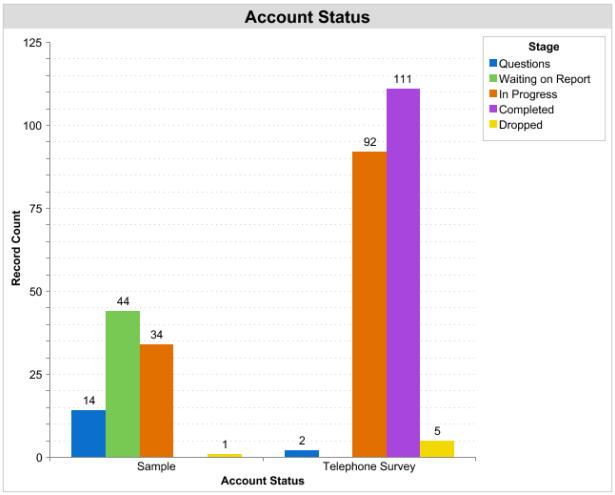
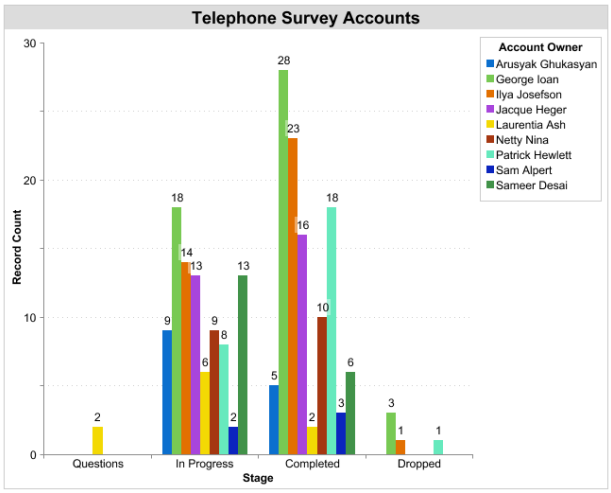


Figure 7. Status of Projects Assigned to Engineers



Because the web-based tracking system is updated continuously, project managers were able to present current information that could be compiled almost instantly; there was no lag in compiling and reporting each evaluator's progress.

Reports were also generated to compile survey results. Mass data exports were used to calculate the MAR. The tracking system data was left unaltered after the MAR survey stage ended, and therefore serves as an unchanging record of how site contacts responded to survey questions.

Appendix G
Early Replacement Survey Instrument

Early Replacement Survey Instrument

The table below shows the survey instrument that was used to gather the responses for the present a summary of the raw data that was collected as a part of the early measure replacement survey.

REP - EARLY REPLACEMENT VERSUS RETROFIT			
REP1_a	REP1_a - Age of replaced REP_A equipment		*To the best of your recollection, how old was the equipment you replaced with **[REP_A]?**[Do you know what year it was originally installed?]*
		01 = [Record Age (Years)]	[RECORD AGE (YEARS)]
		02 = [Record Date (Year)]	[RECORD DATE (YEAR)]
		03 = Did not replace older equipment	DID NOT REPLACE OLDER EQUIPMENT
		94 = Not Yet Installed	NOT YET INSTALLED
		96 = Refused	REFUSED
		97 = Don't Know	DON'T KNOW
		98 = No Answer	NO ANSWER
		99 = Not Applicable	NOT APPLICABLE
REP1_a1	REP1_a1 - Age of replaced REP_A equipment in years	___ Record age in years	___ RECORD AGE IN YEARS
REP1_a2	REP1_a2 - Age of replaced REP_A equipment date given	___ Record date installed	___ RECORD DATE INSTALLED
[IF REP1_a=3, SKIP TO REP2b_a, ELSE READ REP2_a]			
REP2_a	REP2_a - Decision to replace REP_A		Which of the following **best** describes your decision to replace [**REP_A]?**
		01 = It was working but not as effectively as newer models	IT WAS WORKING BUT NOT AS EFFECTIVELY AS NEWER MODELS
		02 = It was working but we needed a larger/smaller system	IT WAS WORKING BUT WE NEEDED A LARGER/SMALLER SYSTEM
		03 = Working but old and would need to be replaced eventually	IT WAS WORKING BUT OLD AND WOULD PROBABLY NEED TO BE REPLACED IN THE NEXT COUPLE OF YEARS ANYWAY
		04 = It required frequent maintenance	IT REQUIRED FREQUENT MAINTENANCE
		05 = It was not working	IT WAS NOT WORKING
		06 = Other (Specify)	OTHER [SPECIFY]
		94 = Not yet installed	NOT YET INSTALLED
		96 = Refused	REFUSED

		97 = Don't Know	DON'T KNOW
		98 = No Answer	NO ANSWER
		99 = Not Applicable	NOT APPLICABLE
REP2_ao	REP2_ao - Verbatim answer for decision to replace REP_A	___ Record verbatim	___ RECORD VERBATIM
[IF REP1_a=3, ASK REP2B_a. ELSE SKIP TO AUD1]			
REP2b_a	REP2b_a - Decision to incorporate REP_A		*Why did you decide to incorporate **[REP_A]?***
		01 = [Record Verbatim]	RECORD VERBATIM
		96 = Refused	REFUSED
		97 = Don't Know	DON'T KNOW
		98 = No Answer	NO ANSWER
		99 = Not Applicable	NOT APPLICABLE
REP2b_ao	REP2b_ao - Verbatim response for decision to incorporate REP_A	___ Record verbatim	___ RECORD VERBATIM

Appendix H
Tabulated Measure Adoption Rates by Year and Category

Table 1. Tabulated MAR Results, 1 of 2

Time from Study Completion to Implementation	Measure Adoption Rate (% of Recommended Savings That Was Implemented)							
	ALL (n=2,452)	Generation (n=34)	Generation Except 2 Large Sites (n=32)	Non-Generation (n=2,418)	Down-state (n=344)	Upstate (n=2,108)	With +kWh Savings (n=1,893)	Without +kWh Savings (n=559)
Prior to study completion	0.01	0.00	0.00	0.01	0.00	0.02	0.01	0.01
Within 1 year	0.25	0.00	0.00	0.34	0.50	0.13	0.26	0.18
1 - 2 yrs	0.34	0.02	0.04	0.46	0.54	0.24	0.35	0.27
2 - 3 yrs	0.46	0.27	0.52	0.54	0.62	0.38	0.48	0.31
3 - 4 yrs	0.61	0.60	0.52	0.61	0.64	0.59	0.63	0.41
4 - 5 yrs	0.62	0.60	50.2	0.62	0.65	0.60	0.64	0.42
5 - 6 yrs	0.65	0.72	0.52	0.63	0.75	0.60	0.68	0.43
6 - 7 yrs	0.65	0.72	0.52	0.63	0.75	0.61	0.68	0.43
7+ yrs	0.65	0.72	0.52	0.63	0.75	0.61	0.68	0.43
Total Recomm (excluding Don't Know)	14,837,762	4,016,360	2,070,192	10,821,402	4,854,254	9,983,508	13,060,163	1,777,599

Table 2. Tabulated MAR Results, 2 of 2

Time from Study Completion to Implementation	Measure Adoption Rate (% of Recommended Savings That was Implemented)							
	Controls (n=598)	Envelope (n=159)	Hot Water (n=62)	HVAC (n=596)	Industrial (n=120)	Lighting (n=409)	Motors (n=171)	Other (n=287)
Prior to study completion	0.00	0.01	0.02	0.03	0.00	0.01	0.01	0.00
Within 1 year	0.64	0.06	0.07	0.23	0.19	0.17	0.11	0.10
1 - 2 yrs	0.73	0.13	0.14	0.33	0.33	0.32	0.30	0.32
2 - 3 yrs	0.78	0.16	0.23	0.40	0.42	0.48	0.31	0.45
3 - 4 yrs	0.84	0.18	0.45	0.47	0.57	0.58	0.36	0.49
4 - 5 yrs	0.85	0.18	0.45	0.50	0.57	0.60	0.36	0.49
5 - 6 yrs	0.85	0.20	0.45	0.51	0.59	0.60	0.36	0.50
6 - 7 yrs	0.86	0.20	0.45	0.51	0.59	0.60	0.36	0.51
7+ yrs	0.86	0.20	0.45	0.51	0.59	0.60	0.36	0.51
Total Recomm (excluding Don't Know)	3,830,474	464,715	109,734	2,940,477	730,795	1,302,957	483,185	910,507

Appendix I
Early Replacement Survey Data

Early Replacement Survey Data

The tables below present a summary of the raw data that was collected as a part of the three survey questions that were a part of the early measure replacement survey.

Summary including 'blank' responses.

REP 1: To the best of your recollection how old was the equipment you replaced?										
	Under 10 years	**10 to 15 years**	**15 to 20 years**	**20 to 25 years**	**Over 25 years**	**Did not replace older equipment**	**Not Yet Installed**	**Don't Know**	**Not Applicable**	**Totals**
Building Envelope	0	0	0	2	3	2	0	0	0	7
Compressed Air Leaks	2	0	1	1	2	0	1	1	0	8
Compressed Air: Other	0	0	4	0	0	2	0	1	0	7
Compressed Air: VFD Compressors	0	1	1	2	0	0	0	0	0	4
Controls	4	1	1	1	2	4	0	4	0	17
Heat Recovery	2	0	0	0	0	1	1	0	0	4
HVAC equipment replacement	2	1	1	1	14	1	0	2	0	22
Insulation	1	1	0	0	0	0	0	2	0	4
Lighting and Lighting Controls	12	3	2	4	9	1	0	2	0	33
Motors replacement	1	1	1	1	4	0	0	0	0	8
Not Applicable	0	0	0	0	0	0	0	0	84	84
Other	2	3	2	1	8	2	0	5	0	23
VFDs	1	0	0	0	4	1	0	1	0	7
WWTP Measures	1	2	0	0	2	0	1	1	0	7
Totals	**28**	**13**	**13**	**13**	**48**	**14**	**3**	**19**	**84**	**235**

Summary excluding 'blank' responses

REP 1: To the best of your recollection how old was the equipment you replaced?										
	Under 10 years	**10 to 15 years**	**15 to 20 years**	**20 to 25 years**	**Over 25 years**	**Did not replace older equipment**	**Not Yet Installed**	**Don't Know**	**Not Applicable**	**Totals**
Building Envelope	0	0	0	2	3	2	0	0	0	7
Compressed Air Leaks	2	0	1	1	2	0	1	1	0	8
Compressed Air: Other	0	0	4	0	0	2	0	1	0	7
Compressed Air: VFD Compressors	0	1	1	2	0	0	0	0	0	4
Controls	4	1	1	1	2	4	0	4	0	17
Heat Recovery	2	0	0	0	0	1	1	0	0	4
HVAC equipment replacement	2	1	1	1	14	1	0	2	0	22
Insulation	1	1	0	0	0	0	0	2	0	4
Lighting and Lighting Controls	12	3	2	4	9	1	0	2	0	33
Motors replacement	1	1	1	1	4	0	0	0	0	8
Other	2	3	2	1	8	2	0	5	0	23
VFDs	1	0	0	0	4	1	0	1	0	7
WWTP Measures	1	2	0	0	2	0	1	1	0	7
Totals	**28**	**13**	**13**	**13**	**48**	**14**	**3**	**19**	**0**	**151**

Percentage of each measure under the total

REP 1: To the best of your recollection how old was the equipment you replaced?										
	Under 10 years	**10 to 15 years**	**15 to 20 years**	**20 to 25 years**	**Over 25 years**	**Did not replace older equipment**	**Not Yet Installed**	**Don't Know**	**Not Applicable**	**Totals**
Building Envelope	0%	0%	0%	29%	43%	29%	0%	0%	0%	100%
Compressed Air Leaks	25%	0%	13%	13%	25%	0%	13%	13%	0%	100%
Compressed Air: Other	0%	0%	57%	0%	0%	29%	0%	14%	0%	100%
Compressed Air: VFD Compressors	0%	25%	25%	50%	0%	0%	0%	0%	0%	100%
Controls	24%	6%	6%	6%	12%	24%	0%	24%	0%	100%
Heat Recovery	50%	0%	0%	0%	0%	25%	25%	0%	0%	100%
HVAC equipment replacement	9%	5%	5%	5%	64%	5%	0%	9%	0%	100%
Insulation	25%	25%	0%	0%	0%	0%	0%	50%	0%	100%
Lighting and Lighting Controls	36%	9%	6%	12%	27%	3%	0%	6%	0%	100%
Motors replacement	13%	13%	13%	13%	50%	0%	0%	0%	0%	100%
Other	9%	13%	9%	4%	35%	9%	0%	22%	0%	100%
VFDs	14%	0%	0%	0%	57%	14%	0%	14%	0%	100%
WWTP Measures	14%	29%	0%	0%	29%	0%	14%	14%	0%	100%
Totals	**19%**	**9%**	**9%**	**9%**	**32%**	**9%**	**2%**	**13%**	**0%**	**100%**

Summary including ‘blank’ responses.

REP 2A: Which of the following *best* describes your decision to replace the equipment?												
	It was working, but not as effectively as newer models	It was working, but we needed a smaller/ larger system	Working, but old and would need to be replaced in the next couple of years anyway	It required frequent maintenance	It was not working	Other	Not yet installed	Refused	Don't Know	No Answer	Not Applicable	Totals
Bldg Envlp	2	0	0	0	3	0	0	0	0	0	2	7
Comp. Air Leaks	3	0	1	2	1	0	0	0	0	1	0	8
Comp. Air: Other	4	0	0	0	0	0	0	0	1	0	2	7
VFD Comp.	3	0	1	0	0	0	0	0	0	0	0	4
Controls	5	0	2	1	1	0	0	0	4	0	4	17
Heat Recovery	2	0	0	0	0	0	0	0	0	1	1	4
HVAC eqpt replace	11	1	6	2	1	0	0	0	0	0	1	22
Insulation	3	0	0	0	0	0	0	0	1	0	0	4
Lighting /Ltg Ctrols	24	1	5	0	0	0	0	0	2	0	1	33
Motors replace	4	0	2	2	0	0	0	0	0	0	0	8
Other	0	0	0	0	0	0	0	0	0	0	84	84
VFDs	3	2	10	1	1	1	0	0	3	0	2	23
WWTP Measures	2	1	3	0	0	0	0	0	0	0	1	7
Totals	4	1	0	0	0	1	0	0	1	0	0	7
Bldg Envlp	70	6	30	8	7	2	0	0	12	2	98	235

Summary excluding 'blank' responses

REP 2A: Which of the following *best* describes your decision to replace the equipment?												
	It was working, but not as effectively as newer models	It was working, but we needed a smaller/ larger system	Working, but old and would need to be replaced in the next couple of years anyway	It required frequent maintenance	It was not working	Other	Not yet installed	Refused	Don't Know	No Answer	Not Applicable	Totals
Bldg Envlp	2	0	0	0	3	0	0	0	0	0	2	7
Comp. Air Leaks	3	0	1	2	1	0	0	0	0	1	0	8
Comp. Air: Other	4	0	0	0	0	0	0	0	1	0	2	7
VFD Comp.	3	0	1	0	0	0	0	0	0	0	0	4
Controls	5	0	2	1	1	0	0	0	4	0	4	17
Heat Recovery	2	0	0	0	0	0	0	0	0	1	1	4
HVAC eqpt replace	11	1	6	2	1	0	0	0	0	0	1	22
Insulation	3	0	0	0	0	0	0	0	1	0	0	4
Lighting /Ltg Ctrols	24	1	5	0	0	0	0	0	2	0	1	33
Motors replace	4	0	2	2	0	0	0	0	0	0	0	8
Other	3	2	10	1	1	1	0	0	3	0	2	23
VFDs	2	1	3	0	0	0	0	0	0	0	1	7
WWTP Measures	4	1	0	0	0	1	0	0	1	0	0	7
Totals	70	6	30	8	7	2	0	0	12	2	14	151

Percentage of each measure under the total

REP 2A: Which of the following *best* describes your decision to replace the equipment?												
	It was working, but not as effectively as newer models	It was working, but we needed a smaller/ larger system	Working, but old and would need to be replaced in the next couple of years anyway	It required frequent maintenance	It was not working	Other	Not yet installed	Refused	Don't Know	No Answer	Not Applicable	Total
Bldg Envlp	29%	0%	0%	0%	43%	0%	0%	0%	0%	0%	29%	100%
Comp. Air Leaks	38%	0%	13%	25%	13%	0%	0%	0%	0%	13%	0%	100%
Comp. Air: Other	57%	0%	0%	0%	0%	0%	0%	0%	14%	0%	29%	100%
VFD Comp.	75%	0%	25%	0%	0%	0%	0%	0%	0%	0%	0%	100%
Controls	29%	0%	12%	6%	6%	0%	0%	0%	24%	0%	24%	100%
Heat Recovery	50%	0%	0%	0%	0%	0%	0%	0%	0%	25%	25%	100%
HVAC eqpt replace	50%	5%	27%	9%	5%	0%	0%	0%	0%	0%	5%	100%
Insulation	75%	0%	0%	0%	0%	0%	0%	0%	25%	0%	0%	100%
Lighting /Ltg Ctrols	73%	3%	15%	0%	0%	0%	0%	0%	6%	0%	3%	100%
Motors replace	50%	0%	25%	25%	0%	0%	0%	0%	0%	0%	0%	100%
Other	13%	9%	43%	4%	4%	4%	0%	0%	13%	0%	9%	100%
VFDs	29%	14%	43%	0%	0%	0%	0%	0%	0%	0%	14%	100%
WWTP Measures	57%	14%	0%	0%	0%	14%	0%	0%	14%	0%	0%	100%
Totals	**46%**	**4%**	**20%**	**5%**	**5%**	**1%**	**0%**	**0%**	**8%**	**1%**	**9%**	100%

Summary including 'blank' responses

REP 2B: Why did you decide to incorporate the replacement measure?						
	Explanation offered (verbatim recorded)	**Refused**	**Don't know**	**No answer**	**Not applicable**	**Totals**
Building Envelope	2	0	0	0	5	7
Compressed Air Leaks	0	0	0	0	8	8
Compressed Air: Other	1	0	1	0	5	7
Compressed Air: VFD Compressors	0	0	0	0	4	4
Controls	4	0	0	0	13	17
Heat Recovery	1	0	0	0	3	4
HVAC equipment replacement	1	0	0	0	21	22
Insulation	0	0	0	0	4	4
Lighting and Lighting Controls	1	0	0	0	32	33
Motors replacement	0	0	0	0	8	8
Not Applicable	0	0	0	0	84	84
Other	2	0	0	0	21	23
VFDs	1	0	0	0	6	7
WWTP Measures	0	0	0	0	7	7
Totals	**13**	**0**	**1**	**0**	**221**	**235**

Summary excluding 'blank' responses

REP 2B: Why did you decide to incorporate the replacement measure?						
	Explanation offered (verbatim recorded)	**Refused**	**Don't know**	**No answer**	**Not applicable**	**Totals**
Building Envelope	2	0	0	0	5	7
Compressed Air Leaks	0	0	0	0	8	8
Compressed Air: Other	1	0	1	0	5	7
Compressed Air: VFD Compressors	0	0	0	0	4	4
Controls	4	0	0	0	13	17
Heat Recovery	1	0	0	0	3	4
HVAC equipment replacement	1	0	0	0	21	22
Insulation	0	0	0	0	4	4
Lighting and Lighting Controls	1	0	0	0	32	33
Motors replacement	0	0	0	0	8	8
Other	2	0	0	0	21	23
VFDs	1	0	0	0	6	7
WWTP Measures	0	0	0	0	7	7
Totals	13	0	1	0	137	151

Percentage of each measure under the total

REP 2B: Why did you decide to incorporate the replacement measure?						
	Explanation offered (verbatim recorded)	**Refused**	**Don't know**	**No answer**	**Not applicable**	**Totals**
Building Envelope	29%	0%	0%	0%	71%	100%
Compressed Air Leaks	0%	0%	0%	0%	100%	100%
Compressed Air: Other	14%	0%	14%	0%	71%	100%
Compressed Air: VFD Compressors	0%	0%	0%	0%	100%	100%
Controls	24%	0%	0%	0%	76%	100%
Heat Recovery	25%	0%	0%	0%	75%	100%
HVAC equipment replacement	5%	0%	0%	0%	95%	100%
Insulation	0%	0%	0%	0%	100%	100%
Lighting and Lighting Controls	3%	0%	0%	0%	97%	100%
Motors replacement	0%	0%	0%	0%	100%	100%
Other	9%	0%	0%	0%	91%	100%
VFDs	14%	0%	0%	0%	86%	100%
WWTP Measures	0%	0%	0%	0%	100%	100%
Totals	**9%**	**0%**	**1%**	**0%**	**91%**	**100%**

Appendix J
Glossary of Acronyms and Definitions

Section 5: GLOSSARY OF ACRONYMS AND DEFINITIONS[55]

AAPOR (American Association for Public Opinion Research) – A leading association of public opinion and survey research professionals.

Construct Validity – The extent to which an operating variable/instrument accurately taps an underlying concept/hypothesis, properly measuring an abstract quality or idea.

Contact Rate – This is one of the final disposition and outcome rates for surveys defined by the American Association for Public Opinion Research (AAPOR).[56] The contact rate includes all outcomes where an eligible respondent was reached and the interview attempted divided by these plus those not contacted. The three contact rate outcomes are completes, refusals, and break-offs (the numerator of the contact rate).

Cooperation Rate – This is one of the final disposition and outcome rates for surveys defined by the American Association for Public Opinion Research (AAPOR).[57] The proportion of all cases interviewed of all eligible units ever contacted. Those contacted (the denominator) includes completes, refusals and break-offs.[58]

FR (Free Riders, Free Ridership) - A program participant who would have implemented the program measure or practice in the absence of the program.

NTG, NTGR (Net-to-Gross, Net-to-Gross Ratio) – The relationship between net energy and/or demand savings, where net is measured as what would have without the program, what would have occurred naturally, and gross savings (often evaluated savings). The NTGR is the ratio of net savings to gross savings. For NYSERDA programs the NTGR is defined as 1 minus free ridership plus spillover (1 – FR + SO).

Refusal Rate – One of the final disposition and outcome rates for surveys defined by the American Association for Public Opinion Research (AAPOR).[59] The proportion of all cases in which an eligible respondent refuses to be interviewed, or breaks-off an interview, of all potentially eligible cases.

Response Rate – One of the final disposition and outcome rates for surveys defined by the American Association for Public Opinion Research (AAPOR).[60] The response rate estimates the fraction of all eligible working numbers where a request for an interview was made. The denominator of this ratio is inclusion of all possible components where a request for an interview could be attempted. More specifically the response rate is the number of completed interviews divided by the sum of: completes, refusals, break-offs, not contacted, and the figure estimated for unknown eligibility.

---

55 Much of this glossary is taken from the 2004 California Evaluation Framework, which was prepared for the California Public Utilities Commission and the Project Advisory Group in September 2004 by a team led by TecMarket Works and including one of the authors of this report from Megdal & Associates.

56 American Association for Public Opinion Research (AAPOR) 2011. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys,* 2011. Each of the rates presented here has multiple, more specific categories and definitions provided by AAPOR. Standard Definitions is available on AAPOR website: www.aapor.org

57 Ibid.

58 Ibid.

59 Ibid.

60 Ibid.

Response rate = (Completes)/(Completes+refusals+break-offs+not contacted+(e*(unknown eligibility)).

SO (Spillover)- Includes **Participant Inside Spillover (ISO)** and **Participant Outside Spillover (OSO)** and **Non-Participant Spillover** -- Reductions in energy consumption and/or demand caused by the presence of the energy efficiency program, beyond program related gross savings of participants.

Inside spillover occurs when, due to the project, additional actions are taken to reduce energy use at the same site, but these actions are not included as program savings.

Outside spillover occurs when an actor participating in the program initiates additional actions that reduce energy use at other sites that are not participating in the program.

Non-participant spillover is the reductions in energy consumption and/or demand from measures installed and actions taken or encouraged by non-participating vendors or contractors because of the influence of the program.

SRE (Stratified Ratio Estimator) – An efficient sampling design combining stratified sample design with a ratio estimator. It's most advantageous when the population has a large coefficient of variation. (A large coefficient of variation occurs, for example, when a substantial portion of the projects have small savings and a small number of projects have very large savings estimates.) The ratio estimator uses supporting information for each unit of the population when this information is highly correlated with the desired estimate to be derived from the evaluation, such as the tracking savings and the evaluated savings.

MAR – Measure adoption rate. The study-level MAR represents the percent of recommended measure savings in completed studies that have been installed. The MAR was determined using a telephone survey.

SRR – Savings realization rate. The SRR is the ratio of the field verified energy savings to the phone verified (MAR) energy savings. The SRR represents the percent of study-estimated savings for implemented measures that the evaluation team estimates as being actually achieved based on the results of the on-site survey and analysis.

Appendix K
MAR Round 2 Memo

Megdal & Associates

M E M O R A N D U M

To: Rebecca Reed, NYSERDA Energy Analysis and the Evaluation Staff of the New York Department of Public Service (DPS)

From: Satyen Moray & Jon Maxwell, ERS; Lori Lewis, Megdal & Associates, LLC; and Kathryn Parlin, WHEC

Subject: Round 2 MAR Results

Date: December 21, 2011

The purpose of this memo is to provide an update to the Flexible Technical Assistance (FlexTech) program's Measure Adoption Rate (MAR) estimates based on the second round (Round 2) of telephone surveys. The first round (Round 1) of MAR telephone surveys was conducted from May 2010 to July 2010, and Round 2 was fielded a year later, allowing additional time for the adoption of measures associated with the FlexTech studies completed in the later years (2008/2009). This memo provides revised MAR estimates developed by combining the data from the first and second rounds of MAR telephone surveys.

The first section describes the overall process that was used to screen the customers for the second round of MAR telephone surveys. The section that follows it describes the approach used to update the MAR estimates as well as the results. The final section summarizes the findings.

Selection Criteria for the Round 2 MAR Studies

The eligibility criteria for Round 2 were based on the time elapsed since the participant received the FlexTech report and on the results of Round 1. The initial sample frame for the Round 2 MAR surveys was composed of the 301 studies with completed surveys from Round 1. This group was further restricted by the criteria discussed below.

- Studies with reports prior to 2004 were excluded from Round 2. The 2004 studies would have been in the 7+ year bin for which no further resolution was likely. The Impact Evaluation Team in consultation with NYSERDA evaluation staff concluded that since the MAR curve plateaus after year 5, the new round of telephone calls have a very small chance of updating the MAR estimates for the studies conducted during this time period.
- Studies from Round 1 were included only if they had any "recommended" (R) or "recommended mutually exclusive" (RME) measures that were not resolved during the Round 1 MAR survey. A single unresolved measure qualified a study for the Round 2 MAR calls. Measures were considered to be resolved if they were identified through the Round 1 as installed or the participant indicated that there was no chance of future implementation.
- Participants who requested not to be contacted again were dropped from Round 2.

Using the above criteria, a total of 117 studies out of the 301 from the Round 1 MAR Survey were eligible for the Round 2, representing 39% of the studies with completed surveys in Round 1. These 117 studies had a total of 1,182 measures, and inquiries were made on 561 unresolved measures. The implementation status of the remaining 621 measures in these studies were completely resolved during Round 1.

MAR Survey Protocols

The Round 1 MAR telephone survey protocols were also used for the Round 2. The key requirement was that each site in the sample had to be contacted a minimum of six times at different times of the day and days of the week. Attempts were made to contact all 117 facilities in the sample. Twenty-three sites were still unresponsive after six or more calls and were dropped as a result. The overall response rate for Round 2 was 80%, which is better than the 71% response rate observed during the more extensive Round 1 MAR telephone survey. There were 237 measures associated with the dropped studies.

MAR Evaluation Approach

This section repeats the MAR algorithm discussed in the main evaluation report. The study-level MAR represents the percent of savings of measures recommended in completed studies that have been installed. The Round 2 MAR was calculated as shown in Equation 1. This is the same equation that was used for the Round 1 MAR calculation.

$$MAR_n = \frac{kWh_{Installed\_in\_Year\_n}}{kWh_{\mathrm{Re}commended}}$$

where,

n = Years between study completion and installation

MAR = Measure adoption rate

$kWh_{Installed}$ = Total kWh savings estimated in the FlexTech studies for all customer-reported installations in year n

$kWh_{Recommended}$ = Total kWh savings for all measures recommended for implementation in the FlexTech studies

A MAR of 1 would indicate that customers had installed all recommended measures in their FlexTech studies.
The MAR estimate was tabulated in a matrix format as shown below:

Year Bin	0	1	2	3	n+
Measure Adoption Rate	MAR0	MAR1	MAR2	MAR3	MARn

where,

MAR0 = MAR for year 0, meaning the first 12 months following study completion

MAR1 = MAR for year 1

MAR2 = MAR for year 2, etc.

Year X is the difference between the year in which the study was delivered to the customer and the year in which it was actually installed. It is likely that a study delivered to the customer in 2008 will have some measures in year bin 0, some in year bin 1, and some in year bin 2 depending on when they were actually installed. This method captures the rate at which the measures are adopted over time.

Units of kWh are presented in the MAR equation above. In the case of non-electric (fuel) savings, the same equation was used, with MMBtus replacing the electric kWh. These results can be disaggregated by upstate/downstate region, energy efficiency vs. on-site generation and measure technology category.

The final MAR for each reporting cycle was the total installed kWh savings for that entire reporting cycle divided by the total recommended kWh savings for that reporting cycle.

The MAR telephone survey followed standard survey methodological procedures:

3. 1. After consulting with NYSERDA, the Impact Evaluation Team modified the Round 1MAR survey instrument slightly to expedite the Round 2 MAR survey by removing questions that were not relevant to the Round 2 MAR survey. Examples of prior questions that were removed from the Round 1 MAR instrument were the funding source and the decision-maker. Since the funding source question yielded successful responses for only less than half the measures during Round 1, it did not add value and was not used in the analysis of the survey results. The decision-maker question was fielded during Round 1 and the required information was already available in the database, so the question did not require repeating.

4. 2. The interviewers called at least six times at different times of the day, different days of the week, and over multiple weeks.

5. 3. The surveys were tailored so that the interviewers did not inquire about measures that were already resolved during Round 1.

Several respondents chose to respond through email exchanges. Overall the response rate exceeded 80%, and interviewers used the tracking system to enter responses.

Inspections and in-person interviews during the SRR site visits early in the year revealed that a material proportion of the MARs reported in the telephone survey were incorrect. The final step in MAR calculation was to use the MAR correction factor that was based on the observed / telephone interview-based adoption rates in the prior SRR sample.

Round 2 MAR Results

MAR Over Time

Figures 1 and 2 illustrate Round 2 and Round 1 MAR curves, respectively, over time. The dashed line represents the percentage of recommended savings adopted each year after study completion; the solid line depicts the cumulative percentage adopted. Eventually slightly over 65% of the savings associated with recommended measures was implemented by study recipients.

Figure 1. Round 2 - Program MAR Over Time

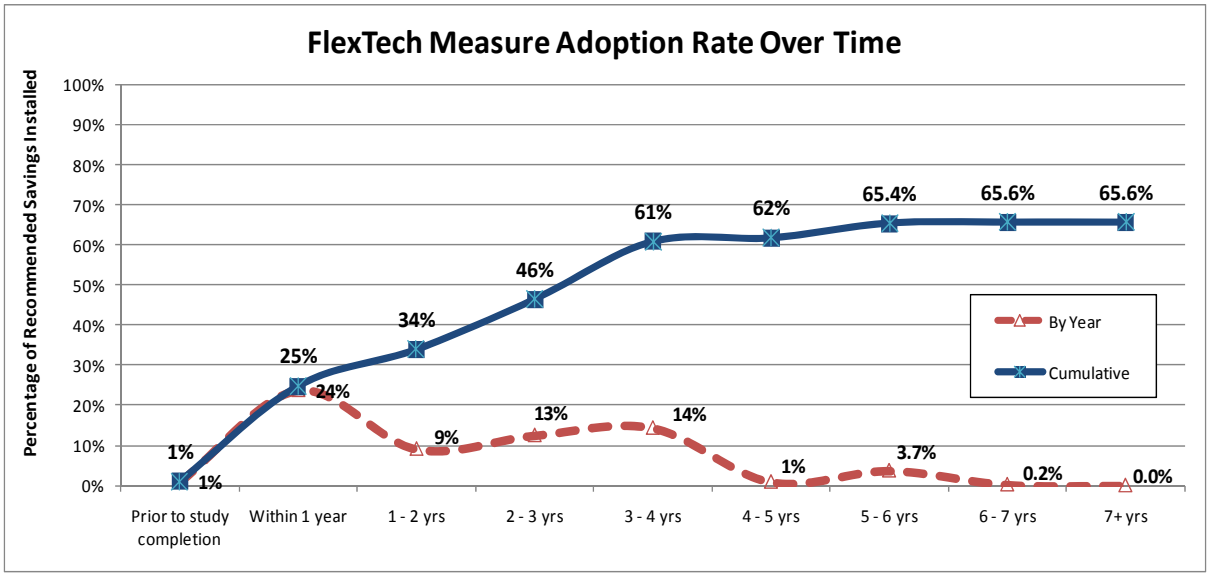
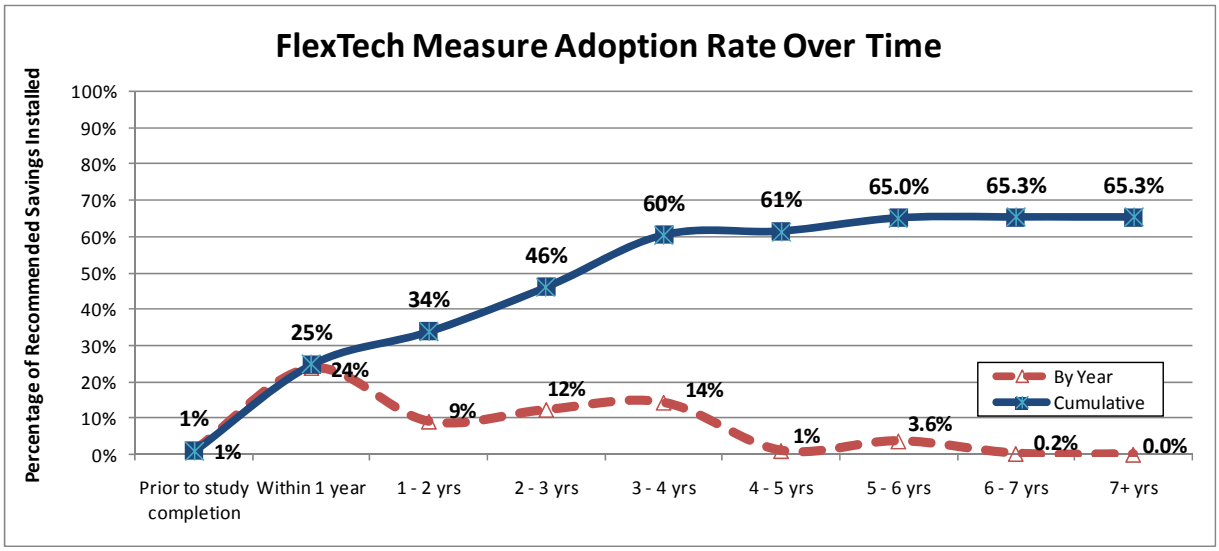


Figure 2. Round 1 - Program MAR Over Time



Comparison of the Round 2 and Round 1 MAR results indicate a marginal increase in the adoption rates for studies in the 3 - 5 year bins. The Round 2 MAR values for the FlexTech studies in the 3 - 5 year bins showed a 1% increase in MAR compared to the Round 1 MAR results

MAR Curve – Generation and Non-Generation

Figures 3 and 4 present the Round 2 and Round 1 MAR curve for the cogeneration and non-cogeneration types.

Figure 3. Round 2 - MAR for Generation and Non-Generation Technologies

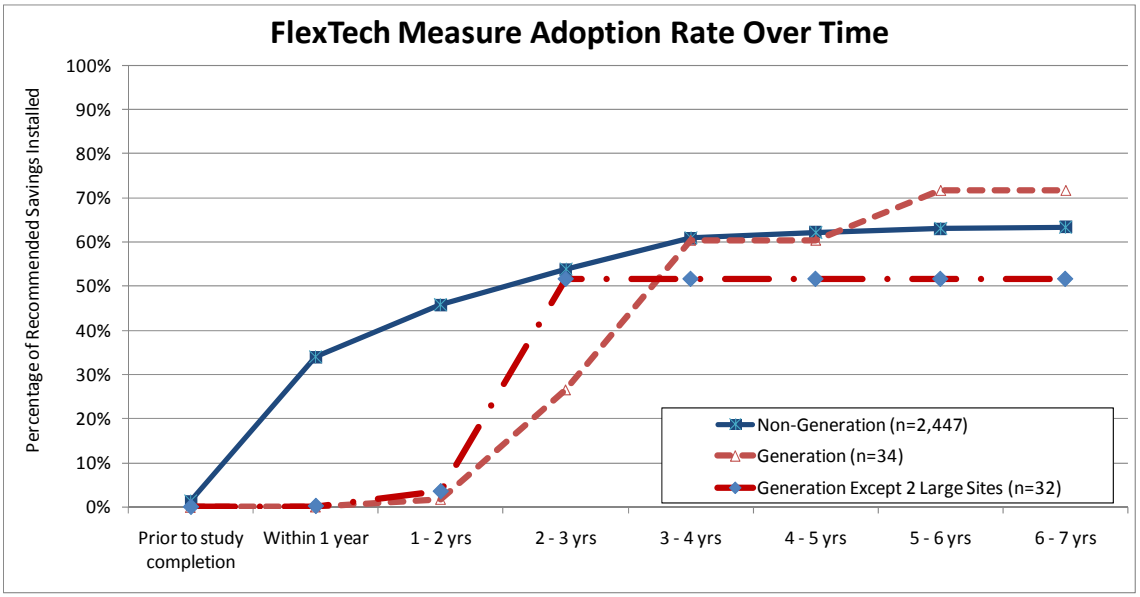
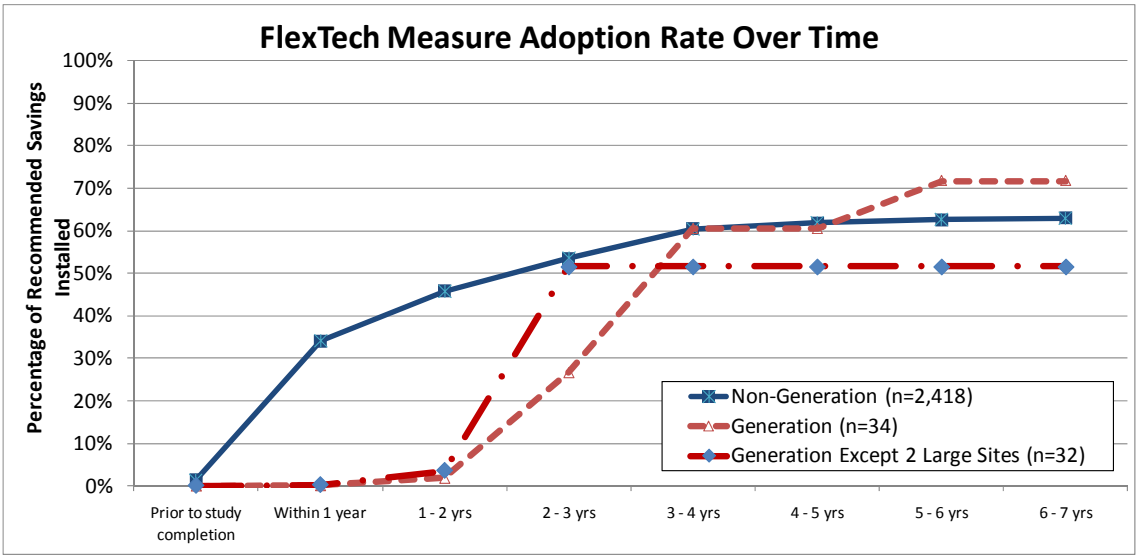


Figure 4. Round 1 – MAR for Generation and Non-Generation Technologies



Round 2 did not result in any change to the cogeneration MAR profile that was developed from Round 1. The Round 2 MAR values for the non-cogeneration projects increased by 0.5% on average for all studies in year bins 1 and above.

MAR Curve – Technology Groups

Figures 5 and 6 present the Round 2 and Round 1 MAR curves, by the technology groups identified in the NYSERDA database for non-generation measures.

Figure 5. Round 2 - MAR Curves for Type of Non-Generation Technologies

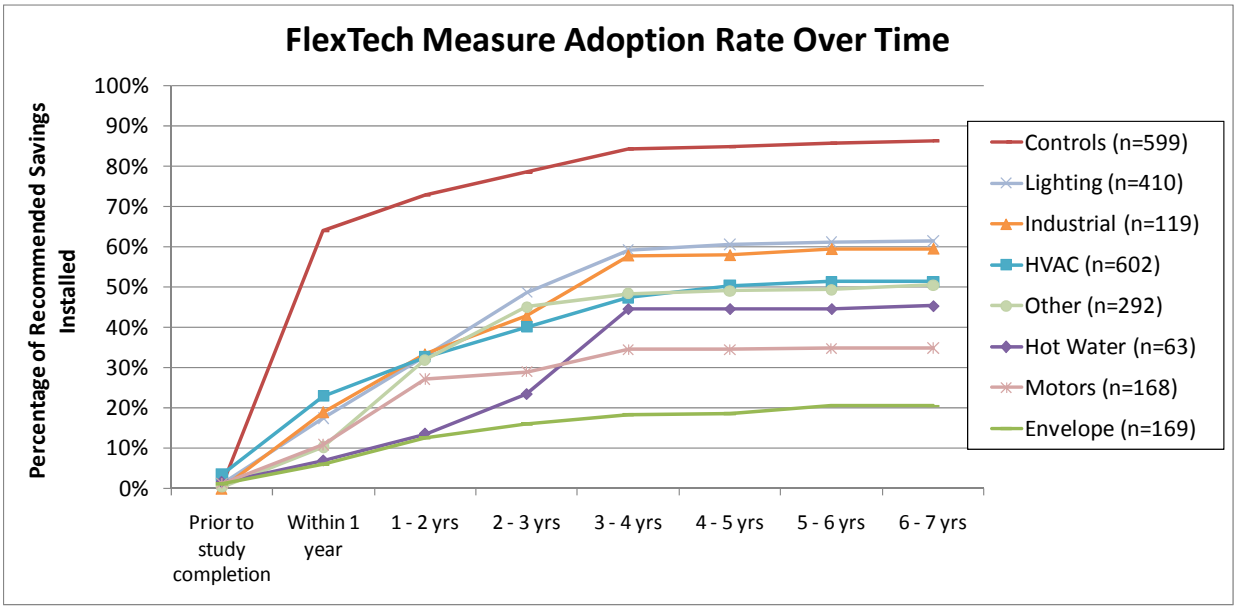
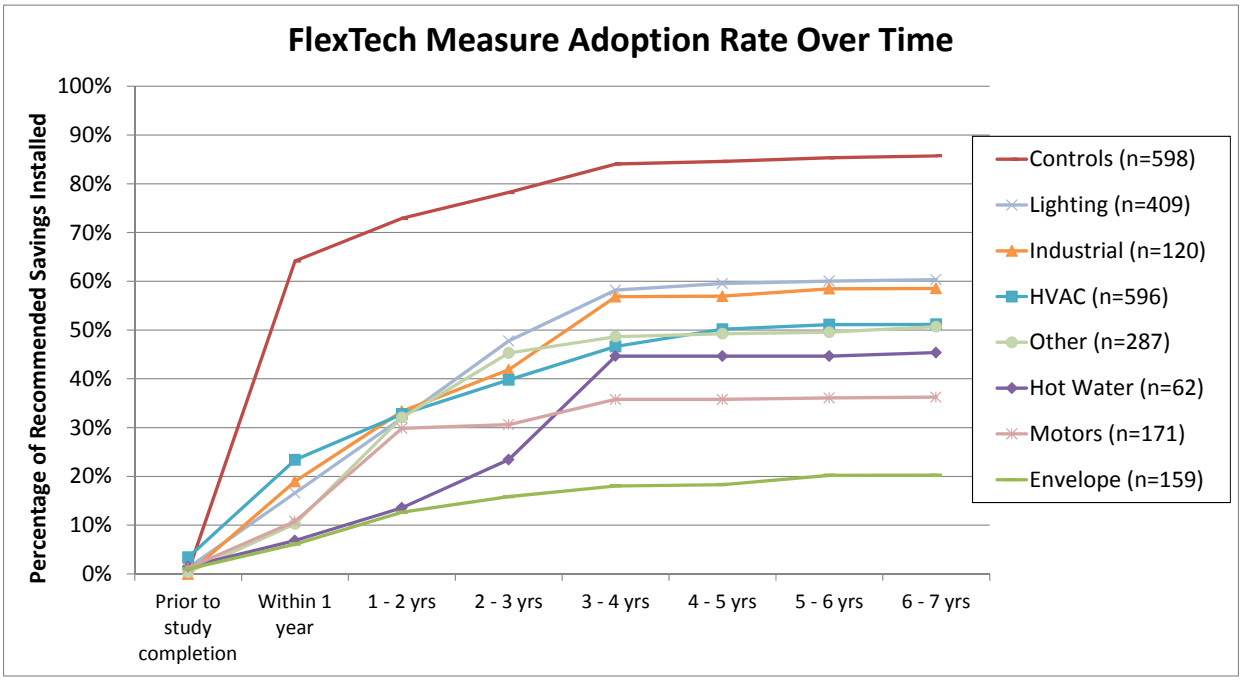


Figure 6. Round 1 – MAR Curves for Type of Non-Generation Technologies



The two plots show that the controls measures are by far the most frequently adopted measure type by study recipients. Envelope measures are the least adopted.

The controls, industrial, and lighting measure groups showed the greatest change in the Round 2 MAR values when compared with the Round 1 MAR values. Controls measures showed an average increase of 0.5% for year bins 2 and greater, industrial measures showed an average increase of 1% for year bins 2 and greater, and lighting showed an average increase of 1.1% for year bins 2 and greater when compared with the Round 1 MAR values. The other technology groups also showed minor increases across the year bins.

MAR Curve for Upstate and Downstate Study Measures

Figures 7 and 8 show the Round 2 and Round 1 MAR curves, respectively, for upstate and downstate measures.

Figure 7. Round 2 - MAR for Downstate and Upstate Measures

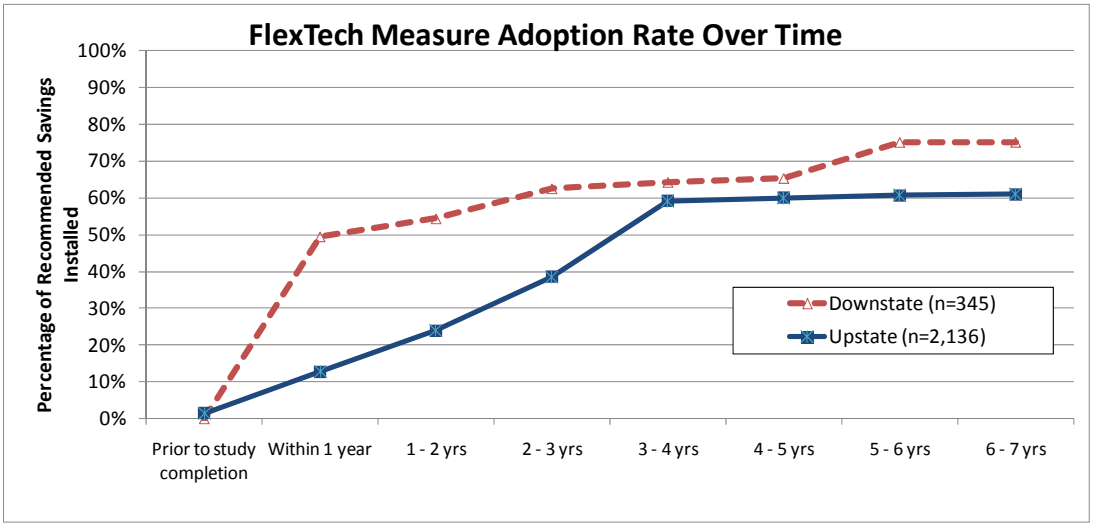
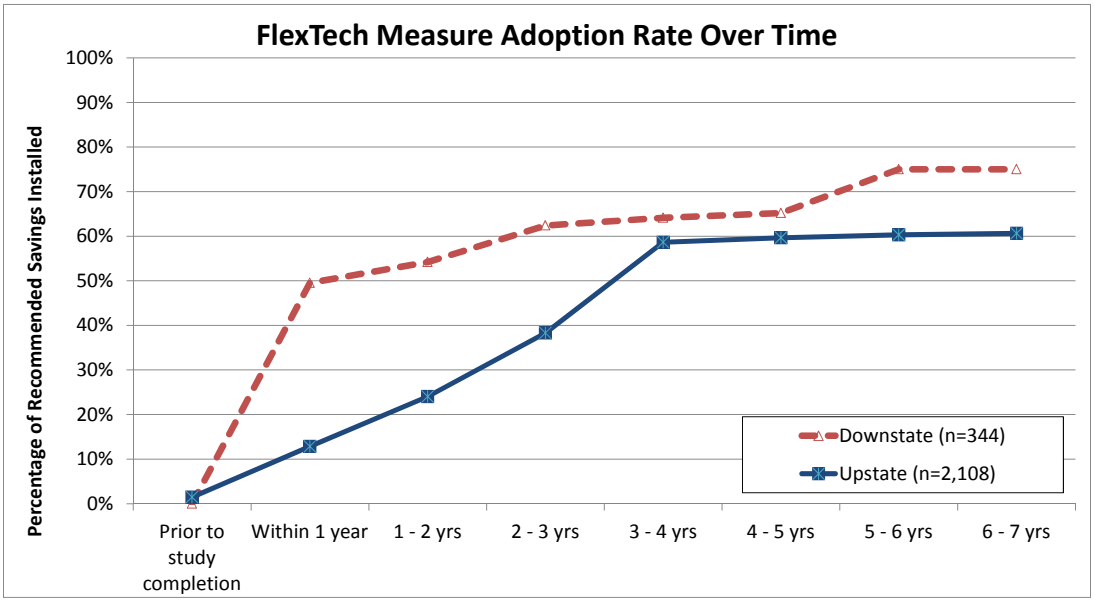


Figure 8. Round 1 - MAR for Downstate and Upstate Measures



In the two plots, the adoption rates differ significantly for upstate and downstate measures in the first few years but gradually converge before another modest departure. The Round 2 MAR values for the downstate measures remained fairly unchanged compared to the Round 1 MAR values, while the upstate measures registered an average MAR increase of 0.4% compared to the Round 1 MAR values for studies in the year bins 2 and greater.

MAR Curve for Measures with and without Electric Savings

Figures 9 and 10 show the Round 2 and Round 1 adoption rates, respectively, for measures that have electric energy savings and those that do not.

Figure 9. Round 2 - MAR for Measures with and without Electricity Savings

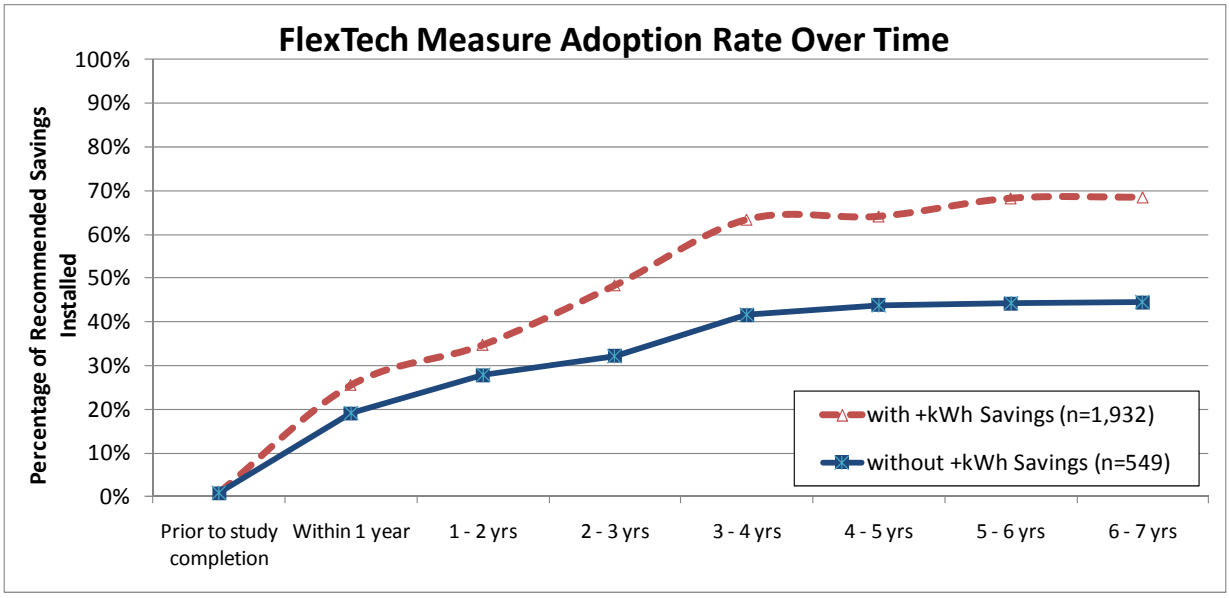
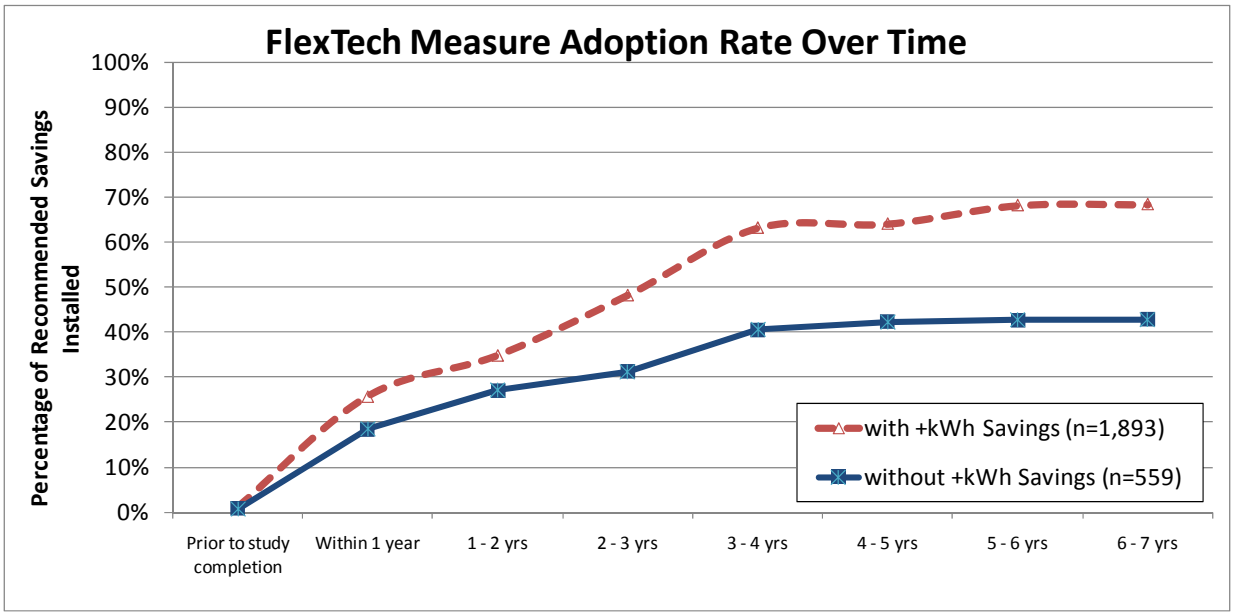


Figure 10. Round 1 – MAR for Measures with and without Electricity Savings



The Round 2 MAR values for the projects with electric energy savings were marginally greater (0.1%) in all year bins compared to the Round 1 MAR values. The Round 2 MAR values for measures that had negative or no electric savings showed a greater increase compared to the Round 1 MAR results. The Round 2 MAR values for the measures with negative or no electric savings values were on average 1.2% greater for all year bins (except the prior-to study bin) compared to the Round 1 MAR values.

Summary of Results

Table 1 summarizes the projected long-term eventual MARs for electric and non-electric efficiency measures and on-site generation based on the combined MAR survey results from Rounds 1 and 2. Non-electric measures have the lowest adoption rate, electric efficiency is just above the 0.66 mean, and generation measures are adopted at a slightly higher rate than efficiency. The overall MAR increased by 0.01 percent compared to the Round 1 MAR results.

Table 1. FlexTech Projected Long-Term Measure Adoption Rates

Measure Type	Long-Term Projected Measure Adoption Rate
Electric Energy Efficiency	0.67
Non-Electric Energy Efficiency	0.42
On-Site Generation	0.72
Overall	0.66

Conclusions

A total of 117 studies out of the 301 from the Round 1 were eligible for Round 2, representing 39% of the studies completed in Round 1. The 117 studies selected for the Round 2 had a total of 1,182 measures out of which inquiries were made on 561 unresolved measures. Of the 117 sites, 23 were dropped as a result of unresponsiveness of the site contacts. These 23 sites had 237 measures. The overall response rate for Round 2 was 80%, compared to 71% for the Round 1 MAR telephone survey. The higher response rate was primarily due to the fact that the surveyors had spoken with the contacts during the prior MAR calls.

The Round 1 MAR results were based on the status of 2,452 recommended measures. Round 2 updated the adoption data for 29 measures that were unresolved in Round 1, while updating the data on 114 other partially implemented measures identified in Round 1. After the completion of the Round 2 calls, 181 measures were still left unresolved (not implemented completely but likely to be implemented in the future).

- Overall the MAR results did not change substantially between the Round 1 and Round 2 MAR surveys. The final combined MAR value increased by 0.01% compared to the Round 1 MAR estimate.
- By measure type:
  - The Round 2 MAR values for the non-cogeneration projects increased on an average by 0.5% for all studies.
  - Industrial and lighting measure groups showed the greatest change in the Round 2 MAR values when compared with the Round 1 MAR values. Industrial measures showed an average increase of 1% for year bins 2 and greater, and lighting measures showed an average MAR value increase of 1.1% for year bins 2 and greater compared to the Round 1 MAR values. The other technology groups also showed minor increases across the year bins.

- The Round 2 MAR values for the downstate measures remained fairly unchanged compared to the Round 1 MAR values while the upstate measures registered an average MAR increase of 0.4% compared to the Round 1 MAR values for studies in the year bins 2 and greater.
- The Round 2 MAR values for the projects with electric energy savings were marginally greater (0.1%) in all year bins compared to the Round 1 MAR values.
- The Round 2 MAR values for measures that had negative or no electric savings showed a greater increase of 1.2% compared to the Round 1 MAR results indicating higher adoption for fossil fuel savings-based measures in the last year.

We do not recommend repeating this survey within the next 12 months, as the results are unlikely to change significantly.