

NYSERDA STATEWIDE RESIDENTIAL POINT-OF-SALE LIGHTING (CFL EXPANSION) PROGRAM

Final Evaluation, Measurement, and Verification Plan

June 15, 2009

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

The goals of the CFL Expansion Fast Track Program evaluation are to assess 1) the extent to which program support engenders more sales than are directly supported by the program; 2) how much remaining potential there is for replacing incandescent bulbs with CFLs; 3) how much net savings are being achieved by the program; 4) the sustainability of program savings; and 5) how to modify the current and future program to maximize cost-effective net savings from efficient residential lighting. Specific plans to evaluate these goals are described in more detail below.

The CFL Expansion Fast Track evaluation budget is tentatively projected to be approximately \$1.1 million through April of 2010; after that time the need for additional evaluation through December of 2011 will be examined. This represents approximately 12% of program costs in 2009, and 5% of program costs in 2010, and is based on the program's expected contribution of electric savings toward NYSERDA's fast track portfolio (42% of the total projected Fast Track portfolio savings, representing substantial risk if the target is not met), along with market dynamics—the extremely rapid pace of change in the CFL market, both in New York and nationally. Because of this rapid pace of change, the first year of the evaluation will be critical to understanding to what extent the residential lighting market has already been transformed, and how much cost-effective net savings have been achieved.

A key aspect of the evaluation activities through early 2010 (including the process evaluation and the 2010 RDD surveys) will be to recommend whether to continue the program as planned with minor modifications, to alter the program significantly and focus on different lighting technologies and specific market segments, or to discontinue it altogether.¹ If the recommendation is to continue the program largely in its current form, then frequent evaluation updates will be necessary because the market is changing so rapidly, and the opportunities for cost effective savings could decline precipitously. If the recommendation is to change the focus of the program, then the structure of the evaluation will likely change in later years, although some of the key elements could remain in place. If the recommendation is to discontinue the program, it could mean ending or sharply curtailing the evaluation after the first-year budget of \$1.1 million has been spent. In any case, along with a recommendation about the future direction of the program, the team will also make a concomitant recommendation about the future direction of the evaluation after the first year.

A schedule for the recommended first-year evaluation activities can be found in Table 10. Data collection for the program baseline study, specifically random-digit-dial (RDD) telephone surveys, was completed in January 2009, and the on-site saturation surveys will be completed in July 2009. A second RDD survey, to evaluate the effects of the 2009 CFL Expansion Program, will be conducted in early 2010. Process evaluation is also planned for early in the program cycle, with completion of the first process evaluation study expected in 2009, with a small addition in early 2010. Depending on the findings and recommendations emerging from the 2009 and 2010 evaluation, data collected through these early efforts will also be updated periodically through future studies of the market.

¹ A determination regarding the specific timing of the 2010 RDD surveys will be made in Fall 2009.

This plan expands on the brief plan outline described in NYSERDA's 60-day filing, which was developed by NYSERDA in conjunction with evaluation contractors and subcontractors Nexus Market Research, Summit Blue Consulting, Lori Megdal, Research Into Action, and APPRISE. This expanded plan was developed by NYSERDA in conjunction with evaluation contractors and subcontractors Nexus Market Research, Summit Blue Consulting, Research Into Action, APPRISE, GDS Associates, and Energy and Resource Solutions.²

II. Program Goals

The **New York Energy SmartSM Products Program** partners with retailers and manufacturers to increase the supply of and demand for energy efficient ENERGY STAR[®] products within NYSERDA's service territory. The Program currently partners with more than 930 retailers and 34 manufacturers of various energy efficient ENERGY STAR products, including CFLs. Current program efforts include working with retailers by providing staff training, point-of-purchase materials for stores, financial assistance with marketing and promotional efforts, and market share incentives. The CFL Expansion Program will: increase marketing and co-op advertising promotions with retail stores and lighting manufacturers; continue to increase the network of retail partners and manufacturers; increase consumer accessibility to a wider variety of CFLs by providing incentives to retailers to increase the number of CFLs sold and increase permanent shelf space for these products; increase in-store promotions and point-of-purchase information to educate consumers; increase participation in the CFL Collection Center Program; and promote the manufacture, sale, and usage of high power factor CFLs.

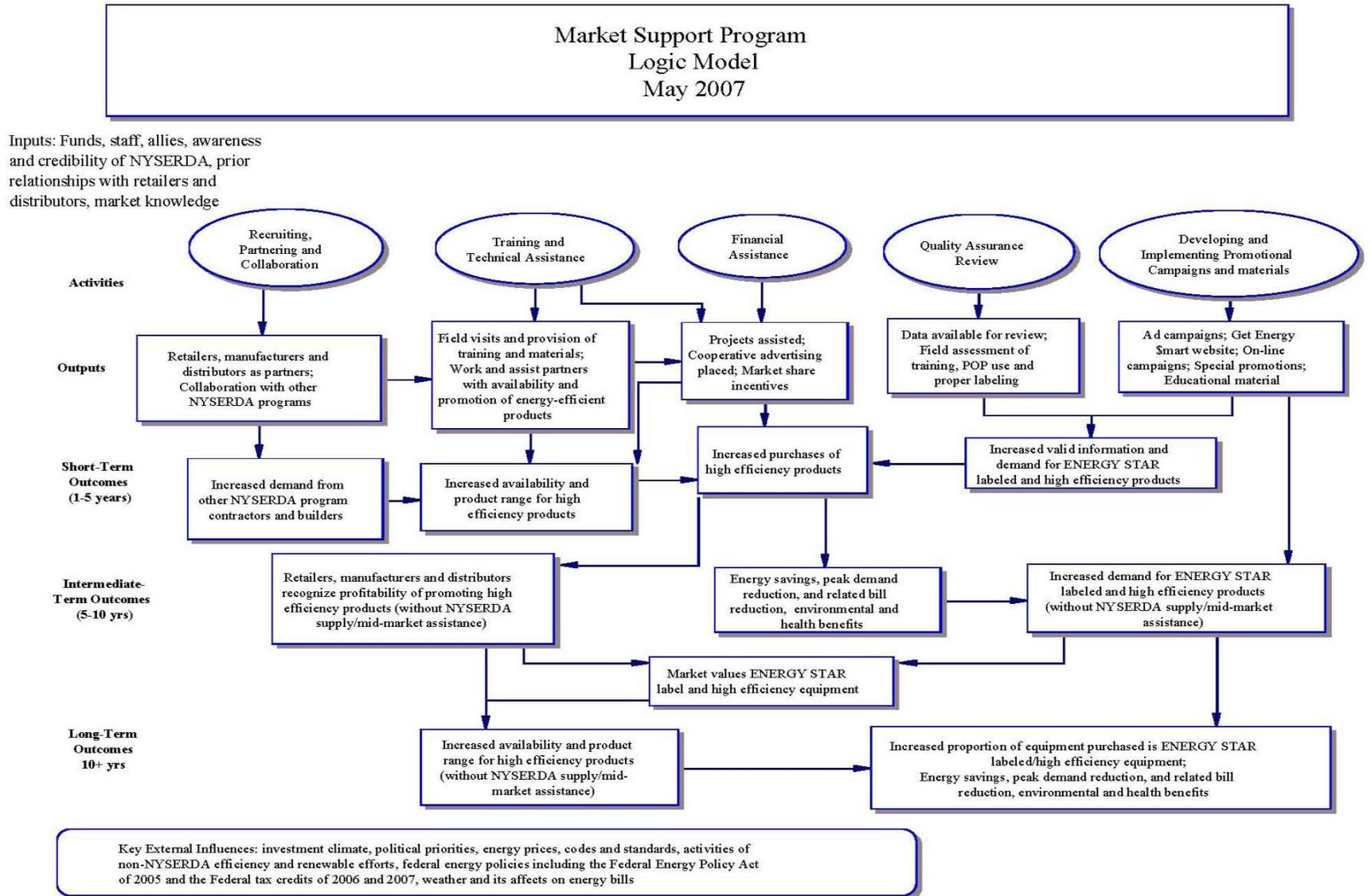
Through the CFL Expansion Program, it is estimated that 16.9 million CFLs will be installed within SBC territory, saving approximately 1,083,861 MWh. Of the estimated 16.9 million CFLs sales the program will impact, 37% (approximately 6,775,701 CFLs) will be purchased as a result of direct program incentives and 63% (approximately 10,151,815 CFLs) of the bulbs will be purchased as spillover from program marketing, outreach, and education efforts

III. Logic Model/Theory

Figure 1 presents NYSERDA's most recent logic model of the Market Support Program, which includes the **New York Energy SmartSM Products Program**. Early in the planned evaluation, NYSERDA will develop a logic model to reflect the CFL Expansion Program.

² The preceding version of the CFL Expansion Evaluation Plan (dated February 20, 2009) called for an evaluation of retail stores in New York State, New York City and the comparison areas described herein to triangulate baseline information gathered through the RDD and saturation studies. The retail surveys were proposed to collect data on metrics such as product placement, shelving practices, and also included interviews with retail store managers. However, due to the delay in receiving approval on the CFL Expansion Evaluation Plan and concerns expressed by the Department of Public Service Staff regarding the overall cost of the evaluation, the NYSERDA evaluation team recommends removing the retailer survey component from the Plan given the length of time that has passed to compare findings to the RDD and since retail store measurements conducted now may not adequately represent baseline conditions. As a result, the NYSERDA evaluation team now supports an alternative approach described in detail in this plan. NYSERDA received approval to conduct the New York State, New York City and comparison area RDD surveys in December 2008.

Figure 1 – Market Support Program Logic Model



While a logic model will be developed to reflect the CFL market specifically rather than ENERGY STAR products generally, some key performance indicators can be identified, as shown in Table 1.

Table 1 – Preliminary Performance Indicators and Sources

Indicator	Source
Number of retail partners recruited	NYSERDA program records
Number of manufacturer partners recruited	NYSERDA program records
Number of CFLs covered in markdown agreements	NYSERDA program records
Number of program CFLs sold	NYSERDA program records
Number of field visits to partner stores	NYSERDA program records
Dollar value of cooperative advertising	NYSERDA program records
Consumer familiarity with CFLs	Random-digit dial survey
Consumer satisfaction with CFLs	Random-digit dial survey
Proportion of households currently using CFLs	Random-digit dial survey, saturation survey
Mean number of CFLs in homes of current users	Saturation survey
Number of CFLs in storage	Saturation survey
Number of CFLs in use	Saturation survey
Saturation of CFLs out of all eligible sockets	Saturation survey
Number of households purchasing CFLs in past three months, past year	Random-digit dial survey
Market-level sales of CFLs	Random-digit dial survey
Market share of CFLs out of all bulbs	Random-digit dial survey
Extent to which manufacturers would continue to develop new products in absence of program	Manufacturer interviews
Extent to which manufacturers are developing and marketing new lighting technology that could affect the residential market	Manufacturer interviews

IV. Evaluation Goals

As noted above, the primary goals of this evaluation effort are to assess: 1) the extent to which program support engenders more sales than are directly supported by the program; 2) how much remaining potential there is for replacing incandescent bulbs with CFLs; 3) how much net savings are being achieved by the program; 4) the sustainability of program savings; 5) how to modify the current and future program to maximize cost-effective net savings from efficient residential lighting, or—if such modifications appear not to be possible—whether NYSERDA should consider ending or greatly curtailing the program. Specific evaluation tasks to meet these goals are described below. NYSERDA expects that the large majority (approximately 90%) of the CFL Expansion Program’s evaluation budget will be allocated to the impact and market evaluation work described herein.

The overarching statewide and out-of-State comparison studies proposed in this evaluation plan are necessary to support a robust evaluation effort of the CFL Expansion Program. In this case, these studies are so integral to the Program’s design and impact evaluation that the evaluation plan includes the full cost of all necessary overarching studies. However, NYSERDA recognizes that other program administrators supporting CFLs could significantly benefit from these study results, in terms of both program implementation and evaluation. Other NY program administrators’ evaluation efforts could also be significantly informed by results of the proposed CFL saturation studies described in this plan. Therefore, NYSERDA respectfully requests that the Department of Public Service (DPS) and the Evaluation Advisory Group (EAG) consider the potential for joint funding for the studies outlined herein. The timeline for completing this initial evaluation and fielding the early baseline study effort did not permit for consideration of this study effort by the newly formed EAG subcommittee on energy efficiency studies

prior to fielding the data collection efforts described. However, should DPS ultimately require other program administrators to jointly fund the CFL baseline work, NYSERDA could offer to hold detailed discussions on the evaluation activities (i.e., CFL saturation study) and consider adding any issues important to other program administrators that aren't already addressed. If this requirement is not made, and joint funding does not occur, the study will be designed as outlined in this Plan to meet the needs of evaluating NYSERDA's program, and an opportunity for additional data collection to benefit other program administrators may be lost.

Note that the research outlined in this plan primarily addresses the residential market.

V. Market Characterization & Assessment and Impact Evaluation Plan

As noted above, the CFL Expansion Fast Track Program is a market-level program, as is the NYSERDA market transformation program it is based upon, the **New York Energy \$martSM** Products Program. These types of programs require market level studies specifically designed to simultaneously measure market characteristics and behavior, and derive estimates of impacts from the market perspective. NYSERDA plans to conduct a combined market and impact evaluation effort for this program.

a. Research objectives

The NYSERDA CFL Expansion Program is taking place in the context of a rapidly expanding national CFL market, with large quantities of CFLs being sold in areas without CFL incentive programs. Moreover, the program is intended to transform the market, and a large portion of the savings is expected to result from spillover. Given this context, the following are the primary market and impact evaluation objectives:

1. To assess the extent to which program support engenders more sales than are directly supported by the program
2. To measure net energy and demand savings attributable to the CFL Expansion Program
3. To estimate how much remaining potential there is for replacing incandescent bulbs with CFLs
4. To assess the effect of the program on consumers' awareness of and attitudes toward CFLs, as well as their CFL usage and purchase patterns
5. To assess the sustainability of program savings
6. To determine whether and in what form further intervention in the residential lighting market is needed, and for how long, in order to continue to produce cost-effective net savings.

b. Research activities

A number of the research tasks in this study will involve comparing New York with other areas that do not have active CFL programs. There are a number of considerations in selecting comparison areas. One such consideration is the fact that this research is occurring at a time when the California Public Utilities Commission is conducting a market effects evaluation of the California investor-owned utilities' CFL programs, which involves a similar research approach involving comparison areas.³ The areas selected for comparison with California—based on the absence of substantial CFL programs, similarity of income and education levels, and the

³ "Compact Fluorescent Lamps Market Effects Scoping Study Findings and Work Plan," prepared for the California Public Service Commission by the Cadmus Group, KEMA, Itron, Nexus Market Research, and A. Goett Consulting, October 31, 2008.

degree of presence of Wal-Mart stores (where many CFLs are sold)—are Pennsylvania, Kansas, and Georgia. It is likely that data from the California study representing these states will be available for NYSERDA to use in its evaluation.⁴⁵ However, these comparison states were selected on the basis of their similarity to California. Therefore, NYSERDA is selecting other comparison areas based on their similarity to New York, and to add to the pool of non-program state data. (Note that it would make sense to make data from this NYSERDA study available to those leading the California evaluation effort). NYSERDA specifically suggests having separate comparison areas for New York City because the types of retailers and the proportion of homeowners vs. renters differ substantially from those in the rest of the State. It should be noted, however, that there is no perfect comparison area—and, in fact, many potential comparison areas are not available because they currently have programs. Even so, with a market transformation program in which net sales (spillover) are projected to be greater than program sales, and program participants are unlikely to be aware of their participation and therefore unable to report on free ridership and spillover, a market-based approach using non-program comparison areas is one of the most powerful analytical tools available—although it may not continue to be an option for much longer because of the proliferation of CFL programs and the dwindling number of non-program comparison areas.

To select comparison areas for the state of New York minus New York City, Nassau County, and Suffolk County,⁶ NYSERDA reviewed the list of non-program states used in the California study less those covered by other studies, and focused on the three states with household incomes closest to that of New York: Delaware, North Carolina, and Ohio. Then, NYSERDA reviewed selected demographic characteristics and big box store counts in each of these three states along with New York (minus New York City, Nassau County, and Suffolk County). As shown in Table 2, it appears that Delaware and Ohio are closer to New York than is North Carolina on most characteristics. For Delaware, these characteristics include percentage of foreign-born persons, the percentage of people speaking a language other than English at home, and per capita income. For Ohio, these characteristics include population growth rate, percentage of white persons, home ownership rate, housing units in multi-unit structures, persons below the poverty level, and concentration of big box stores. One key difference between Ohio and Delaware is that the former is fairly large in area, while the latter is very small, providing residents easy access to stores outside its borders. While having only one comparison area would leave the study more vulnerable (i.e., that one area could institute a CFL program, leaving NYSERDA without a non-program comparison area), NYSERDA will very likely have access to data from additional comparison states through the California study. NYSERDA therefore recommends using Ohio as its sole comparison area. One part of Ohio—the Cincinnati area, which is served by Duke Power—has a CFL program. NYSERDA therefore proposes excluding that area from its Ohio data collection activities.

⁴ Some of NYSERDA's evaluation contractors are involved in California evaluation.

⁵ As discussed in more detail below, similar efforts are also being conducted in Massachusetts, Michigan, and Wisconsin, plus two comparison states. These states have already committed to the multistate modeling effort.

⁶ Nassau and Suffolk Counties receive electric service from Long Island Power Authority, and thus are not included in the NYSERDA SBC territory.

Table 2 – Characteristics of New York State (minus New York City and Long Island) and Potential Comparison Areas

Characteristic	NY less NYC and LI	Delaware	North Carolina	Ohio	Closest to NY
Population, percent change, April 1, 2000 to July 1, 2006	1.00%	8.92%	10.03%	1.10%	OH
Persons 65 years old and over, percent, 2006	13.81%	13.40%	12.20%	13.30%	DE
White persons, percent, 2006	84.35%	74.50%	74.00%	84.90%	OH
Black persons, percent, 2006	8.14%	20.90%	21.70%	12.00%	OH
Asian persons, percent, 2006	2.62%	2.80%	1.90%	1.50%	DE
Persons of Hispanic or Latino origin, percent, 2006	6.27%	6.30%	6.70%	2.30%	DE
Foreign born persons, percent, 2000	7.31%	5.70%	5.30%	3.00%	DE
Language other than English spoken at home, pct age 5+, 2000	11.60%	9.50%	8.00%	6.10%	DE
Bachelor's degree or higher, pct of persons age 25+, 2000	28.90%	25.00%	22.50%	21.10%	DE
Home ownership rate, 2000	67.12%	72.30%	69.40%	69.10%	OH
Households, 2000	3,118,586	298,763	3,132,013	4,445,773	NC
Housing units in multi-unit structures, 2000	30.2%	18.7%	16.1%	24.1%	OH
Per capita money income, 1999	\$22,378	\$23,305	\$20,307	\$21,003	DE
Persons below poverty, percent, 2004	11.14%	9.60%	13.80%	11.70%	OH
Cost of living—equivalent salary (gi.money.cnn.com/tools/costofliving/costofliving.html)	50,000 (Buffalo) \$51,826 (Plattsburgh)	\$53,653 (Wilmington)	\$48,799 (Charlotte) \$52,400 (Raleigh)	\$50,991 (Cleveland) \$47,546 (Dayton)	NC, OH
Cost of living—equivalent salary (bestplaces.net/COL)	50,000 (Buffalo) \$51,378 (Syracuse)	\$56,224 (Wilmington)	\$58,415 (Charlotte) \$68,112 (Raleigh)	\$50,408 (Cleveland) \$47,101 (Dayton)	OH
Cost of living—equivalent salary (swz.salary.com/costoflivingwizard)	50,000 (Buffalo) \$47,341 (Syracuse)	\$56,315 (Wilmington)	\$55,288 (Charlotte) \$53,651 (Raleigh)	\$52,047 (Cleveland) \$50,618 (Dayton)	OH
Wal-Marts per 100,000 people	1.22	1.29	1.75	1.56	DE
Home Depots per 100,000 people	0.69	1.05	0.49	0.55	OH
Lowes per 100,000 people	0.55	1.05	1.24	0.71	OH
Big box stores per 100,000 people	2.46	3.40	3.48	2.81	OH

Sources: US Census, walmart.com, homedepot.com, lowes.com—accessed November 6, 2008; and gi.money.cnn.com/tools/costofliving/costofliving.html; bestplaces.net/COL; swz.salary.com/costoflivingwizard—accessed December 5, 2008.

To select comparison areas for New York City, NYSERDA first began examining large Northeastern cities without substantial CFL programs. These include Philadelphia, Baltimore, and Washington, DC.⁷ Of these, Washington DC is relatively close to New York City with respect to some key characteristics that might affect CFL purchase patterns, such as the rate of population change, percentage of people over 65, home ownership, housing units in multi-unit structures, cost of living, and concentration of big box stores. NYSERDA added Houston to the list because of its greater similarity to New York City based on characteristics such as percentage of black persons, persons of Hispanic or Latino origin, and people speaking a language other than English at home (see Table 3). The comparison areas likely to be available from other studies (Pennsylvania, Georgia, and Kansas) do not provide good matches to New York City. As mentioned earlier, having only one comparison area would leave the study more vulnerable to the possible loss of its non-program comparison area (i.e., if a program is instituted); NYSERDA therefore recommends using both Washington, DC and Houston as comparison areas.

⁷ Boston was not examined because of an existing program.

Table 3 - Characteristics of New York City and Potential Comparison Areas

Measure	New York City	Philadelphia	District of Columbia	Baltimore	Houston	Closest to NYC
Population, percent change, April 1, 2000 to July 1, 2006	2.57%	-4.56%	1.66%	-3.00%	9.77%	DC
Persons 65 years old and over, percent, 2006*	11.70%	14.10%	12.30%	13.20%	8.40%	DC
White persons, percent, 2000*	44.70%	45.00%	38.40%	31.60%	49.30%	Phil.
Black persons, percent, 2000*	26.60%	43.20%	56.50%	64.30%	25.30%	Hous.
Asian persons, percent, 2000*	9.80%	4.50%	3.20%	1.50%	5.30%	Hous.
Persons of Hispanic or Latino origin, percent, 2000*	27.00%	8.50%	8.20%	1.70%	37.40%	Hous.
Foreign born persons, percent, 2000	35.90%	9.00%	12.90%	4.60%	26.40%	Hous.
Language other than English spoken at home, pct age 5+, 2000	47.60%	17.70%	16.80%	7.80%	41.30%	Hous.
Bachelor's degree or higher, pct of persons age 25+, 2000	27.40%	17.90%	39.10%	19.10%	27.00%	Hous.
Home ownership rate, 2000	30.20%	59.30%	40.80%	50.30%	45.80%	DC
Households, 2000	3,021,588	590,071	248,338	257,996	717,945	Hous.
Housing units in multi-unit structures, 2000	83.2%	21.2%	60.2%	25.8%	37.1%	DC
Per capita money income, 1999	\$22,402	\$16,509	\$28,659	\$16,978	\$20,101	Hous.
Persons below poverty, percent, 1999**	21.20%	22.90%	18.30%	22.90%	19.20%	Phil./ Balt.
Cost of living—equivalent salary (gi.money.cnn.com/tools/costofliving/costofliving.html)	\$50,000 (Queens) \$71,766 (Manhattan)	\$40,268	NA	\$39,468	\$29,097	Phil.
Cost of living—equivalent salary (bestplaces.net/COL)	\$50,000 (Queens) \$51,690 (NYC overall)	\$29,004	\$52,134	28,757	\$26,943	DC
Cost of living—equivalent salary (swz.salary.com/costoflivingwizard)	\$50,000 (Queens) \$62,950 (Manhattan)	\$42,955	\$56,010	\$38,784	\$35,766	DC
Wal-Marts per 100,000 people	0.00	0.28	0.00	0.48	0.75	DC
Home Depots per 100,000 people	0.21	0.41	0.17	0.48	0.75	DC
Lowes per 100,000 people	0.02	0.21	0.00	0.32	0.42	DC
Big box stores per 100,000 people	0.23	0.90	0.17	1.27	1.91	DC

*For DC, year 2006; **For DC, year 2004; Sources: US Census, walmart.com, homedepot.com, lowes.com—accessed November 6, 2008; and gi.money.cnn.com/tools/costofliving/costofliving.html; bestplaces.net/COL; swz.salary.com/costoflivingwizard—accessed December 5, 2008

Specific evaluation tasks are described below.

Net Sales Impact. Measurement of net impact will occur at a market level, focusing initially on net sales rather than net energy and demand savings. This will involve estimating CFL sales in New York (excluding New York City and Long Island) with Ohio as a comparison area, and separately for New York City with the District of Columbia and possibly Houston as comparison areas. The method will rely on the random-digit dial (RDD) survey of consumers, in which consumers will be asked to estimate the number of CFLs they have purchased, with responses verified and possibly calibrated by the on-site saturation results. To mitigate issues of uncertainty such as self-selection bias and unknown dissimilarities between the selected baseline or comparison areas and New York, NYSERDA may also rely on other sources, such as counts of CFL imports from the U.S. Department of Commerce, and sales estimates from the U.S. EPA.

To maximize the analytical potential of the market and impact evaluation of NYSERDA's CFL Expansion Program, NYSERDA suggests a multi-state modeling approach, taking advantage of the research that has been done or is being done in other states. For example, as described above, the California Public Utilities Commission is conducting a market effects evaluation of the California investor-owned utilities' CFL programs, which involves RDD surveys, on-site saturation surveys, and in-store surveys in California and comparison states; a decision by California on joining the multi-state modeling effort is pending. Wisconsin Focus on Energy, Massachusetts program sponsors, and Consumers Energy in Michigan are launching research efforts (all led by NMR) with RDD surveys and on-site saturation surveys (but not in-store surveys) in the three states plus two comparison areas, and have already committed to the multi-state modeling approach. Program sponsors in other states may soon be joining the effort as well.

Should NYSERDA move ahead with the rest of the CFL Expansion Program evaluation work as proposed in this Plan, the sum total of these studies would provide a rich set of data from states and areas with no programs, states and areas with new or recently expanded programs, and states and areas with long-standing programs. Pending approval and participation by NYSERDA in the multi-state modeling effort, the team will conduct analysis of covariance and regression on self-reported purchases, controlling for other factors, thus seeking to explain the effect of the program on sales, and to establish a modeled baseline of CFL sales in New York in the absence of the program.

The approach is to develop a modeled comparison or baseline area taking advantage of variability among households, counties, and states. Some of the program states—such as California, Massachusetts, and Wisconsin—have had long-standing programs varying in their reliance on markdowns instead of coupons, and some of them—such as Maryland (a comparison area for the Massachusetts research), Michigan, and New York—have more recent or newly expanded programs. Other states and areas—Ohio, the District of Columbia, Houston, Indiana (a comparison area for the Wisconsin research) and possibly Georgia, Kansas, and Pennsylvania if California joins the effort—have had no CFL programs until now. Having data from states varying on the continuum of program activity as well as demographic and economic characteristics could help explain the effects of programs—both short-term and long-term—on CFL sales.

Because no comparison area is a perfect match with New York, reflecting what would have happened if NYSERDA had no lighting program, the modeled comparison area approach is superior to the straight single comparison area approach. The “modeled comparison area” would be based on characteristics of households in the samples of each included area, plus characteristics of the county (e.g., number of big box stores per household, foreclosure rate, unemployment rate) and characteristics of the broader area (program dollars spent on CFL incentives, program dollars spent on other CFL promotion, years program in existence). Hence it is not essential to achieve a match between New York and a single other area, but rather to contribute to the variability of household characteristics, general economic characteristics, and program intensity/ history in the overall pool of data.

Saturation. There will also be a separate multi-state model to explain the effects of programs on saturation and number of CFLs installed and in storage, again relying on the random-digit dial (RDD) survey of consumers, with responses verified and possibly calibrated by the on-site saturation results. NYSERDA plans to visit a randomly

selected sample of homes in New York (excluding New York City and Long Island), New York City, and the comparison areas to estimate the number of CFLs out of all eligible sockets that are currently installed, and the number in storage.⁸ NYSERDA will examine the results by room type, since hours of use vary. This will provide an estimate of the remaining potential for CFLs. NYSERDA will collect model numbers of installed CFLs in New York, and ask respondents when and where they were purchased to identify likely markdown CFLs. If there are differences between the sales levels in New York and the comparison areas, the saturation levels may help explain the difference—for example, if sales are lower in New York but saturation is higher, it could mean that earlier sales in New York are displacing current sales. This saturation assessment differs from the current New York energy efficiency potential study being performed by Optimal in that it involves actual onsite inspections of people’s homes and counts of bulbs and sockets rather than reliance on secondary data.

Gross Impact. NYSERDA will estimate the energy and demand savings from program-supported CFLs through the sources identified in Table 4.

Table 4 - Summary of Gross Impact Sources for CFLs

Element	Source
Number of program CFLs sold	Program records
Wattage of program CFLs sold	Program records
Installation rates	Studies from other areas and telephone surveys
Wattage of replaced bulb	Studies from other areas
Hours of use	Studies from other areas
Measure life	Studies from other areas
Demand factor	Studies from other areas

Other recent or soon to be completed studies that could inform this component of the CFL Expansion Program evaluation include the following:

- A lighting logger study in New England providing hours of use, demand factor, wattage of replaced bulb, and installation rates⁹
- A lighting logger study in California providing hours of use, demand factor, wattage of replaced bulb, and installation rates¹⁰
- A measure life study in New England¹¹

⁸ As part of the 2009 RDD data collection, respondents were asked if they would be interested in participating in the on-site saturation study component. The NYSERDA evaluation team retained a list of those interested respondents.

⁹ Nexus Market Research, Inc. and RLW, Inc., *Residential Lighting Markdown Impact Evaluation—Draft*, November 17, 2008, conducted for Cape Light Compact, Connecticut Energy Conservation Management Board, Connecticut Light and Power, Efficiency Vermont, National Grid, NSTAR Electric, United Illuminating, Unitil, and Vermont Department of Public Services. The final report should be completed in the next few weeks.

¹⁰ KEMA, Inc., *Residential Lighting Metering Study: Overview of Study Protocols*, prepared for California Public Utilities Commission, June 4, 2008.

¹¹ Nexus Market Research, Inc. and RLW, Inc., *Residential Lighting Measure Life Study*, June 4, 2008, conducted for Cape Light Compact, Connecticut Energy Conservation Management Board, Connecticut Light and Power, Efficiency Maine, Efficiency Vermont, National Grid, New Hampshire Electric Coop, NSTAR Electric, Public Service of New Hampshire, United Illuminating, Unitil, Vermont Department of Public Services, and Western Massachusetts Electric.

The installation rate will address not only the proportion of recently purchased CFLs that are currently installed, but the rate at which CFLs in storage are put into sockets; this will be derived from responses to questions in the RDD survey, and will be compared to values from other studies.

In calculating lifetime gross savings, NYSERDA will consider the possible effect of a pending Federal law that will restrict and phase out sales of incandescent bulbs. For example, if CFL measure life is estimated to be six years, then the wattage of the replaced bulb for CFLs installed in 2009 may have to be reduced after 2012—or even before that if interviews with manufacturers (see the Process Evaluation section) indicate that they will be phasing out incandescent bulbs before the deadline.

Net Savings Analysis. NYSERDA will apply the net-to-gross sales ratio to the gross savings estimate in order to arrive at net energy and demand savings attributable to the CFL Expansion Program.

Sustainability Assessment. NYSERDA will also examine the sustainability of the market changes associated with the program, assessing the extent to which these changes would continue should program activity be withdrawn or scaled back. There are many other influences on the CFL market, including a Wal-Mart initiative to double the sale of CFLs, promotion of CFLs by the popular press as a strategy for individuals to address climate change, and the recently passed Energy Bill requiring more efficient lighting beginning in 2012.

To assess the extent to which the CFL market has already been transformed, as well as the extent to which these factors will continue to increase CFL sales in the absence of current program efforts, NYSERDA will include a number of additional questions in the manufacturer interviews and has already included relevant questions on satisfaction in the customer surveys discussed earlier in this document. Specifically, the team intends to adopt the approach developed in Massachusetts and devise interview guide/survey questions that address the sustainability questions posed by Hewitt.¹²

¹² Hewitt, D.C. 2000. “The Elements of Sustainability.” In *Efficiency & Sustainability, Proceedings of the 2000 Summer Study on Energy Efficiency in Buildings*. Washington DC: American Council for an Energy-Efficient Economy. Pp. 6.179-6.190.

Table 5 shows how Hewitt’s original questions could be rephrased to fit the CFL market, and also summarizes how the questions might be answered.¹³ As applicable, these questions will be asked in terms of how they pertain to both rebates and to marketing campaigns.

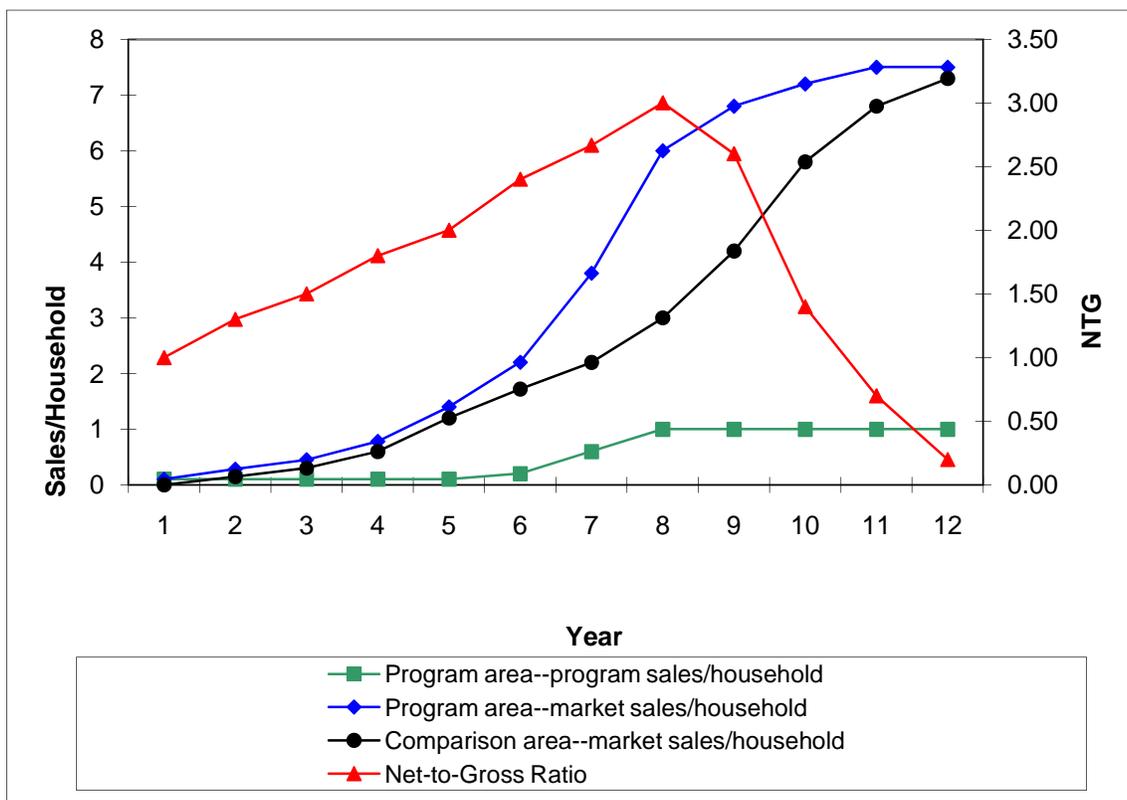
Table 5 - Assessment of Sustainability of the CFL Market

Issue	Response
<i>Will manufacturers continue to develop and market CFLs and will retailers continue to market them without individual regional program support? (Original: Has a private market developed to continue the facilitation?)</i>	Yes. Several manufacturers have announced that they are building new, higher capacity factories in China to accommodate the increased international CFL market demands, and report that CFLs are profitable. The ENERGY STAR program has revised specifications for CFLs, fixtures using the GU-24 technology, and SSL in the form of LEDs, which will become effective during 2008, and manufacturers say they will have products meeting the new specs.
<i>Are CFLs now a mainstream option? (Original: Has the profession or trade adopted it as a standard practice?)</i>	Somewhat. Awareness of CFLs is nearly universal in New York, but consumers still choose incandescents over CFLs for many applications. xx% of upstate New York and xx% of New York City households still do not use CFLs. CFLs are available in a broad range of store types, but drug stores, convenience stores, and discount stores still have limited offerings.
<i>Would it be difficult or costly to revert to earlier equipment—that is, going back to incandescents? (Original: Would it be difficult or costly to revert to earlier equipment or practices?)</i>	Not yet, but relatively soon: Federal legislation EISA 2007 passed in December of 2007 will phase out inefficient light bulbs beginning in 2012.
<i>Are end-users requesting or demanding CFLs? Would there be sufficient consumer demand without regional program support? (Original: Are end-users requesting or demanding it?)</i>	Yes. CFL sales have increased dramatically in areas without program support—2008 sales in the non-program comparison areas of Ohio, the District of Columbia, and Houston, respectively, are xx million and xx million CFLs, or x.x, x.x and x.x per household. However, consumer education is still important for encouraging consumers to use CFLs in more applications and to choose products that will satisfy their lighting needs.
<i>Have the risks to private market actors for manufacturing or marketing CFLs been reduced or removed? (Original: Have the risks to private market actors been reduced or removed?)</i>	Yes. Demand for CFLs nationwide and globally has increased. Many manufacturers announced plans to expand their manufacturing facilities in China this year. Federal EISA 2007 legislation will also encourage the development of more efficient lighting technologies, including incandescents, CFLs, and LEDs. However, the issue of CFL recycling to avoid unsafe mercury disposal still remains unresolved and will become an even more important issue with a greater number of spent CFLs requiring disposal in the coming years.
<i>Are purchasers satisfied with CFLs? (Original: Are purchasers satisfied with it)</i>	Yes. xx% of current CFL users are ‘satisfied’ or ‘very satisfied’ with the products.

¹³ Modified from *Market Progress and Evaluation Report (MPER) for the 2007 Massachusetts ENERGY STAR Lighting Program*, prepared by Nexus Market Research for Cape Light Compact, National Grid, NSTAR, Western Massachusetts Electric, and Unitil, June 13, 2008.

One function of an effective market transformation program is to accelerate the typical S-shaped market adoption curve as posited by diffusion of innovation theory;¹⁴ part of the sustainability assessment will be to examine where CFLs are on the market adoption curve. In the hypothetical, idealized example in Figure 2, in year one the program is responsible for all CFLs sales in the program area market, and the comparison area has no sales; if the program did not exist, there would be no sales in the program area market. From a sales perspective—that is, without considering actual vs. expected savings—the NTG is 1.0. Beginning in the second year, the program starts affecting the program area market, even while the comparison area market is developing to a lesser extent, and the NTG increases through year eight, to a high of 3.0. After that, however, both the program area and the comparison area markets continue developing, with non-program sales catching up as the market becomes transformed; thus the NTG falls below 1.0 by year eleven, and to barely above zero by year 12. The gap between the curves for program area market sales and comparison area market sales—minus program sales—represents the market effects of the program—again, essentially accelerating the market adoption curve and moving it earlier, to the left. In the case of CFLs, the idealized curve is likely to be interrupted—and the “natural” market potential increased to a much higher level—as implementation of the new Federal law restricting the sale of incandescent bulbs approaches.

Figure 2 – Hypothetical Sales and Net-to-Gross Ratio



¹⁴ For a summary of such models, see Geroski, P. A. “Models of technology diffusion,” *Research Policy* 29 (2000) 603 – 625, and Research Triangle Institute, *Market Penetration of New Technologies, Programs, and Services*, Palo Alto, CA: Electric Power Research Institute, 1991.

NYSERDA does not have access to sufficient information either in New York or in its comparison areas to allow the construction of such a graph based on actual historical data, but will have a one-year snapshot upon completion of the baseline study. However, some historic sales information is available from other areas, including program areas (Massachusetts, Wisconsin, the Pacific Northwest) and non-program areas (Michigan), and there will be additional one-year snapshot sales data from a program area (California) and other non-program areas (Pennsylvania, Georgia, Kansas). In addition, the EPA has CFL sales data by state for a limited number of national chains, which could provide a minimum CFL sales figure for each state for 2007 and 2008 (although these data may be more complete for some states than for others, which confounds interpretation). Together, these sources should help determine where the New York market is on the curve depicted in Figure 2. For example, if market-level sales are leveling off in other program areas that have been more active for longer than has NYSERDA's program, it would signal that those markets are approaching their long-term potential, at least for standard 15-25 Watt CFLs. If comparison area sales are comparable to the sales levels of the states with the most active and enduring programs, then it may indicate that the market has been transformed.

However, saturation—the proportion of sockets occupied by CFLs, along with the number of CFLs in storage—is another vital piece of information for assessing the state of the market in New York. For example, if not only the sales but also the saturation levels of the states with the most active and enduring programs are leveling off, and if households in those states have substantial numbers of CFLs in storage that they intend to use to replace existing CFLs, that would further signal the transformation of those markets. However, if the saturation and the number of CFLs in storage are lower in New York than in the most active states—even if sales are just as high or higher—then there is still the possibility of accelerating the market adoption curve until the “steady state” level of saturation and storage is achieved—again, at least for standard 15-25 Watt CFLs; whether such acceleration is later achieved would be measured by comparing saturation and storage rates (as well as sales) over time in New York and the comparison areas. Any such potential would have to be estimated with consideration of the time remaining before implementation of the new federal law restricting the sale of incandescent bulbs, and the likely responses of manufacturers in the interim.

Also note that the process evaluation (see below) will involve an examination of “best practice” residential lighting programs, which will indicate how other program managers view the current and future lighting market, and may provide ideas for the evolution of the NYSERDA program.

If the study determines that the market has not been fully transformed, it will provide feedback on what else must be done, and for how long, to achieve a sustainable change. For example, there may be retail sales channels that warrant additional attention and outreach, and there may be additional “niche” opportunities for future program design (e.g., dimmable and three-way CFLs). The study could also recommend focusing more strongly on LEDs in order to accelerate the market adoption curve of that technology.

Other Market Characterization and Assessment Activities. Consumer telephone surveys will ask about awareness, use, purchases and attitudes toward CFLs, as well as CFL storage, removal, and disposal. Also, the consumer telephone survey will assess awareness and use of LEDs. The telephone survey analysis will also involve segmentation to identify the characteristics of innovators, early adopters, early majority, later majority, and laggards with respect both to CFLs and LEDs—in keeping with the diffusion of innovation model (mentioned earlier) as laid out by Everett Rogers.¹⁵ This approach may help guide program marketing.

NYSERDA also plans to conduct interviews with CFL manufacturers to assess how they view the market; these interviews will be conducted as part of the process evaluation, but will include market-related questions. The questions will address the future direction of the lighting market, including plans for continued manufacturing of

¹⁵ Rogers, Everett M. (2003). *Diffusion of Innovations*, 5th ed.. New York, NY: Free Press.

incandescent bulbs in light of the federally mandated phaseout beginning in 2012, and plans for new technologies such as LEDs.

Research Activities in Future Years. NYSERDA's intention is to carry out the evaluation activities through early 2010 as outlined in this plan and, based on the results, make a recommendation about whether the program should be a) continued more or less as it is at least for the short term, b) altered significantly to focus on different lighting technologies and specific market segments, or c) largely discontinued. The evaluation findings through early 2010 and the resulting recommendation will inform NYSERDA's evaluation approach in later years. If the recommendation is to discontinue the program, the evaluation in later years will entail at most only limited tracking research to assess whether market gains are sustained. If the recommendation is to continue the program, in any form, the current plan tentatively assumes continued use of the comparison area approach. At that point, NYSERDA will assess whether such an approach continues to be viable in light of the rapidly expanding sales of CFLs nationwide, the proliferation of CFL programs and the consequent limited availability of comparison areas, the approaching end of incandescent bulbs in the marketplace because of the new federal law, and the specific form the program takes in later years. If the next generation NYSERDA program emphasizes products that are not promoted by programs in the comparison areas, then the comparison area approach may continue to work. This could be the case even if the promoted products are standard CFLs, although the recommendation may be to focus on specialty CFLs such as three-way and dimmable bulbs, as well as new technologies such as LEDs.

The fact that the potential disappearance of viable comparison areas is an issue points not only to the success of CFLs, but also to the fact that lighting markets—much more so than markets for new homes, for example—are national. In seeming recognition of this fact, a hopeful development in recent residential lighting evaluations has been increased sharing of data and methodologies. A close to ideal solution to NYSERDA's, and possibly other NY program administrators', evaluation needs after the first year would be to become part of a group of state and regional evaluators funding a 48-state (plus District of Columbia) approach, allowing statistical modeling to control for demographics, electric rates, concentration of big box stores, latitude, and program activity, thus seeking to explain the effect of the program on sales, possibly at the level of individual program components, such as prior program activity, level of incentives, types of incentives (e.g., markdown vs. coupon), amount of advertising, etc. Given the number of state markets addressed in evaluations in recent years—California, Georgia, Kansas, Pennsylvania, Iowa, Massachusetts, Maryland, Vermont, Wisconsin, Indiana, Michigan, Oregon, Washington, Idaho, Montana and potentially New York, Ohio, and the District of Columbia (as well, possibly, as the substate market of Houston)—this may not be as much of a stretch as it might seem at first. More modestly, NYSERDA, in cooperation with other sponsors, could draw in additional comparison areas (*i.e.*, beyond the three surveyed in 2009), thereby increasing the number of comparison states even if the total falls short of all 48, and still expand the variability of program and customer characteristics available for statistical modeling. Either way, the intra- and inter-state coordination issues are daunting, and planning for such an effort would have to start immediately in order for this approach to become a possibility for later in 2010.

c. Populations/samples

The New York Department of Public Service suggests that samples be designed to achieve 90/10 precision at the statewide level and at the individual utility level, if possible.

Samples for the RDD survey of households in New York and the comparison areas are shown in Table 6 below.¹⁶ In addition to achieving the 90/10 precision level for individual utilities, another consideration in determining the sample size for the RDD survey—because RDD

¹⁶ As noted above, the NYSERDA evaluation team has completed its 2009 consumer RDD. However, the approach described in this plan also applies to future RDD studies (e.g., 2010).

respondents will be recruited for saturation/socket counts in their homes (described below)—is having sufficient sample sizes for the saturation study. NYSEDA recommends an RDD sample size of 1,000 for New York as a whole (minus New York City, Nassau County, and Suffolk County), and sample sizes of 500 for New York City and each comparison area. (See Table 6.) Because self-reporting on number of installed CFLs has been shown to be very unreliable, NYSEDA instead plans to rely on the saturation surveys to estimate number of CFLs per household. The sample sizes for the saturation survey are 200 in New York State, and 100 for New York City and each of the comparison areas (as shown in

Table 7). A primary driver of the telephone survey sample sizes is using them to recruit households for the saturation survey; for this purpose, a five-to-one ratio seems necessary.

With a sample size of 1,000 for New York as a whole (minus New York City, Nassau County, and Suffolk County), the maximum sampling error (at a 90% confidence level) should be close to +/- 10% for individual utility service territories for purposes of measuring proportions such as awareness of CFLs. Sampling error for the two smallest utilities (Central Hudson Gas & Electric and Orange & Rockland) may be somewhat higher, although they could be lower for indicators with proportions much lower or much higher than 50% of the sample. For example, if the survey indicates that 86% of residents in Central Hudson Gas & Electric territory are aware of CFLs—a level similar that found in other areas in recent surveys—sampling error will be +/-9.8%. As mentioned, the samples for New York City and for each comparison area will be 500. Importantly, for purposes of comparing proportions in New York (minus New York City and Long Island) to those in Ohio or to those in New York City—to determine if there are significant differences— maximum sampling error will be +/-4.5%; for comparing New York City to the District of Columbia or Houston the maximum sampling error will be +/-5.2%. The random-digit dial (RDD) survey in New York State is designed to represent the entire area (New York minus New York City, Nassau County, and Suffolk County) as a whole. As mentioned earlier, one purpose of the RDD survey is to recruit households for the saturation survey. With a sample of 200, NYSEDA does not plan to break out the saturation survey by utility service area. If the RDD sample were designed with disproportionate stratification to increase the precision for smaller utilities, then the precision for the overall sample—both for the RDD survey and the saturation survey—would decrease. NYSEDA believes that, with a proportionate RDD sample of 1000 spread over New York State, the likely sample sizes for the smallest utilities would be adequate.

The formula for computing sampling error of proportions is as follows:

$$\text{Sampling error (90\% confidence)} = 1.645 \sqrt{(p(1-p))/n}$$

The formula for determining whether the difference between two proportions is significant at the 90% confidence level is as follows:

$$(\text{Abs } p_1 - p_2) > 1.645 \sqrt{((p_1(1-p_1))/n_1) + ((p_2(1-p_2))/n_2)}$$

Based on means and standard deviations found in previous surveys, with an RDD sample of 1,000 (New York state minus New York City and Long Island), a mean of 7.5 CFLs purchased would entail a range estimate of 7.0 to 8.0 CFLs at the 90% confidence level, or +/- 7.0%. With a sample of 500 (New York City and each of the comparison areas), a mean of 7.5 CFLs purchased would entail a range estimate of 6.8 to 8.2 CFLs at the 90% confidence level, or +/- 9.9%. With comparisons of means (T-tests) in New York State vs. a comparison area, or New York State vs. New York City, assuming standard deviations similar to those in found in other studies, a difference between means of 7.5 and 6.7 or fewer CFLs purchased would be statistically significant at the 90% confidence level. In comparing means in New York City and a comparison area, a difference between means of 7.5 and 6.5 or fewer CFLs purchased would be statistically significant at the 90% confidence level.

The formula for computing the confidence interval of a mean is as follows:

$$\text{Confidence interval (90\% confidence)} = \text{mean} \pm ((1.645 * \text{sd}) / \sqrt{n})$$

The formula for determining whether the difference between two means is significant at the 90% confidence level is as follows:

$$Abs(mean_1 - mean_2) / \sqrt{((sd_1^2 / n_1) + (sd_2^2 / n_2))} > 1.645$$

The discussion has focused on confidence intervals and sampling errors related to the so-called “Type I Error” in which one concludes a difference exists between two groups when it really does not (i.e., in statistical language, rejects the null hypothesis when it should be accepted). NYSERDA, however, also proposes to test the probability of the so-called “Type II Error” in which one concludes no difference exists, when, in reality, there is a difference (i.e., accepts the null, when it should be rejected).¹⁷ This probability is determined by the *statistical power* of the data, such that the probability of a Type II error is denoted as β and the statistical power is $1 - \beta$ (note that the probability associated with the Type I error is denoted as α and $1 - \alpha$ is the confidence level). One way of improving statistical power and thus reducing the chances of a Type II error is to increase sample size, but this step can actually increase the chance of a Type I error because small differences achieve statistical significance when sample sizes are large. From a practical standpoint, larger sample sizes also increase the costs of conducting a study. Researchers, then, must balance the desire to avoid both Type I and Type II errors within reasonable budgets. Therefore, applied social and behavioral researchers commonly accept 80% as a desirable level of statistical power (or 20% probability of Type II error), which generally allows for reasonable sample sizes that do not lead to overly sensitive tests of statistical significance for Type I errors at the 90% confidence level.¹⁸

Because NYSERDA recommends viewing the results of the RDD and Saturation Studies conducted in 2009 (but focused on products obtained in 2008, prior to the CFL Expansion Program) as baselines, sample sizes were not selected based on the 80% power criterion. NYSERDA is, in essence, assuming that there will be no statistical difference between New York State or New York City and their comparison areas. However, in order to examine this assumption, we will still conduct tests of statistical significance and of statistical power, providing the associated probabilities to aid in interpretation of the findings. Results achieving 80% statistical power will be flagged, as will results achieving 90% statistical significance. For future replications of the RDD and saturation studies, we will use the results of the baseline study together with findings from other studies to develop sample sizes that will allow for an 80% level of statistical power and 90% confidence level. NYSERDA remains hopeful that these revised sample sizes will fall close to those reported for the 2008 and 2009 studies because, as the CFL Expansion Program gains momentum, we expect to see larger programmatic effects, which will manifest at smaller sample sizes while still achieving the desired levels of statistical power and significance. For example, let one assume that the baseline saturation in New York State is 10% (sample size 200) and in Ohio is 6% (sample size 100). This result achieves a low power of 36% (i.e., a 64% probability of Type II error) and does not reach statistical significance at the 90% confidence level.¹⁹ Yet, if we assume that in 2011 that New York State improves to 26% and Ohio to just 11%, even keeping the same sample sizes, statistical power improves to 94% (i.e., a 6% probability of Type II error) and the result achieves statistical significance at the 90% confidence level.

To address possible non-response bias in the RDD survey, NYSERDA made multiple callbacks to maximize response rates. To address selection bias when comparing New York (and New York City) to the various comparison areas, NYSERDA also intends to control statistically for *observed* variables in a regression analysis. To further address selection bias, NYSERDA will also control for *unobserved* variables through the use the inverse Mills ratio when conducting regression analyses.

¹⁷ NYSERDA wishes to acknowledge the thoughtful and careful attention that Richard Ridge paid to the issue of statistical power. His efforts have greatly informed our discussion of it in this work plan.

¹⁸ *HyperStat Online Contents* Chapter 11: Power. 2009. <http://davidmlane.com/hyperstat/A3266.html> Accessed 2/16/09.

¹⁹ Statistical power determined by the statistical program *G-Power 3*. Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. 2007. “G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences.” *Behavior Research Methods*, 39, 175-191.

As mentioned, NYSERDA evaluation contractors will recruit respondents to the RDD surveys for participation in in-home saturation surveys and socket counts. As shown in

Table 7, NYSERDA plans on conducting 200 such in-home visits in New York (minus New York City and Long Island), and 100 in New York City and each of the comparison areas. For purposes of measuring proportions such as number of households with at least one CFL or proportions of sockets filled with CFLs, sampling error will be a maximum of +/- 5.8% for New York (minus New York City and Long Island), and +/-8.3 for New York City and each of the comparison areas. For purposes of comparing New York (minus New York City and Long Island) to Ohio or to New York City—to determine if there are significant differences—maximum sampling error will be +/-10.1%, and for comparing New York City to the District of Columbia or Houston the maximum sampling error will be +/-11.7%, and probably will be lower than +/-10% if the saturation, as expected, is substantially lower than 50%.

In the saturation survey, with a sample of 200 (New York state minus New York City and Long Island), an estimate of 20% saturation (a level found recently in other studies) would entail sampling error of +/-4.7%, and with a sample of 100 (New York City and each of the comparison areas), sampling error would be +/-6.6%. In comparing 20% saturation in New York State (minus New York City and Long Island) to Ohio or New York City, a level of 12% in the latter areas would be significantly different. In comparing 20% saturation in New York City to the District of Columbia or Houston, a level of 11% in the latter areas would be significantly different.

Non-responses bias entails systematic differences between those who do and do not respond to a survey.²⁰ To address possible non-response and self-selection bias in the saturation/socket count survey, NYSERDA will compare the characteristics of households in this survey with those included in the telephone survey and the census. In this analysis, NYSERDA will take advantage of the fact that saturation survey respondents will be a subset of the telephone survey respondents. For example, it may be that telephone survey respondents who did not agree to onsite visits report having fewer CFLs in their homes than do those who agree to the on-sites, perhaps reflecting greater enthusiasm toward CFLs among the latter. In that case, NYSERDA may consider weighting the saturation survey results based on characteristics of the overall telephone survey respondents, or perhaps based on the census. Also, NYSERDA will consider providing incentives of \$100 for residents of New York State (minus New York City and Long Island), Ohio, and (if included) Houston who agree to an on-site saturation survey, and \$150 to residents of New York City and the District of Columbia; the incentive should help mitigate non-response and self-selection bias. If cooperation rates for the saturation survey are high enough, however, then NYSERDA may be able to correct for possible self-selection bias by making sure that those selected for the saturation survey match the overall RDD sample on key characteristics.

Response bias is “an error in given answers to an interviewer’s questions due to misinterpretation by the participant, or the participant responding in such a way that they believe the interviewer wishes them to answer as opposed to their true feelings. Therefore, response bias can occur both deliberately and unintentionally”.²¹ Some potential response bias may be mitigated by the fact that residents of New York and residents of the comparison states will be asked the same questions, although the two groups could differ in some ways that lead to different rates of response bias. Also, some response bias could be identified by comparing what respondents say to what is observed in the on-site visits for saturation surveys; for example, respondents might think it is socially desirable to have CFLs and therefore claim to have more installed than is actually the case, which could be determined (and corrected for) in the on-site visit.

²⁰ <http://www.marketresearchterms.com/n.php>, accessed January 14, 2009.

²¹ <http://www.marketresearchterms.com/r.php>, accessed January 14, 2009.

d. Data collection

The schedules of these data collection efforts are linked: the RDD survey (Table 6) will be used to recruit respondents for the saturation survey (Table 7). Fieldwork for the 2009 RDD survey took place in December of 2008 and January of 2009, while the saturation survey will take place in July of 2009.²² The data collection for the RDD survey—which was budgeted to require an average of 15 minutes to administer—was conducted by Braun Research under the direction of APPRISE, NYSERDA’s data collection contractor. The data collection for the saturation survey in New York State, New York City, and Ohio, will be conducted by individual independent contractors who have experience with in-home assessments as auditors for Weatherization Assistance Program (WAP) agencies; the contracts will be with the individuals, not the WAP agencies, and will be conducted under the direction of APPRISE. In the District of Columbia and Houston, the data collection for the saturation survey will be conducted by GDS Associates and ERS, both of which are members of NYSERDA’s impact evaluation contractor team.

The RDD telephone survey instrument was designed to be comparable with the survey instruments from Massachusetts and California for analysis purposes. The survey instrument was pretested prior to fielding to ensure that all questions were understandable and that all survey skip patterns were properly programmed. The telephone interviewers were trained on the survey questions and survey procedures. During the training, interviewers conducted mock interviews so that they became completely familiar with the survey instrument prior to conducting interviews with respondents. APPRISE research staff monitored interviews throughout the field period to ensure that interviewers were effective in collecting the data and coding the survey responses. In addition, Braun Research conducted on-going monitoring to assess the quality of work by all interviewers. The survey was translated into Spanish for use in the New York City and Houston samples.

In New York State, New York City, and Ohio, the in-home saturation surveys will be conducted by experienced auditors who work with community-based organizations that deliver baseload electric efficiency and weatherization services for the WAP program, as well as for ratepayer funded energy efficiency programs (e.g., NYSERDA EmPower and Ohio EPP). These auditors have extensive experience collecting information from households regarding their use of CFLs and other energy-related behaviors. All auditors selected for the study will be trained by ERS staff regarding the appropriate data collection procedures. These auditors will follow a detailed protocol for collecting CFL information in the selected homes. A sample of the work completed by auditors will be observed by ERS professionals to ensure that the protocols are properly implemented.

²² As noted previously, a determination regarding the specific timing of the 2010 RDD survey will be made in Fall 2009.

Table 6 – CFL Expansion Program Random-Digit Dial (RDD) Survey Specifics

Target Group	Estimated Population Size	Estimated Sample Size	Expected Sampling Precision (90% confidence)*	Data Collection	Survey Administration By	Expected Fielding Dates
Households in National Grid service territory	1,371,754	469	3.8%	RDD	Survey Contractor	Dec. 2008, Early 2010
Households in NYSEG service territory	620,660	212	5.7%	RDD	Survey Contractor	Dec. 2008, Early 2010
Households in O&R service territory	168,720	58	10.9%	RDD	Survey Contractor	Dec. 2008, Early 2010
Households in CHG&E service territory	102,758	35	14.1%	RDD	Survey Contractor	Dec. 2008, Early 2010
Households in RG&E service territory	325,734	111	7.8%	RDD	Survey Contractor	Dec. 2008, Early 2010
Households in Westchester County	337,142	115	7.7%	RDD	Survey Contractor	Dec. 2008, Early 2010
Total NY less NYC & Long Island (NYS)	2,926,768	1,000	2.6%			
Households in Ohio	4,445,773	500	3.7%	RDD	Survey Contractor	Jan. 2009, Early 2010
NYS compared to Ohio or NYC			4.5%			
Households in New York City	3,021,588	500	3.7%	RDD	Survey Contractor	Dec. 2008-Jan. 2009, Early 2010
Households in Washington, DC	248,338	500	3.7%	RDD	Survey Contractor	Jan. 2009, Early 2010
Households in Houston	717,945	500	3.7%	RDD	Survey Contractor	Jan. 2009, Early 2010
NYC compared to DC or Houston			5.2%			

* Proportion of 50%

Table 7 – CFL Expansion Program Saturation Survey Specifics

Target Group	Estimated Population Size	Estimated Sample Size	Expected Sampling Precision (90% confidence)*	Data Collection	Survey Administration By	Expected Fielding Dates
Households in NY less NYC & Long Island (NYS)	2,926,768	200	5.8%	Home Visits	Survey Contractor	July 2009
Households in Ohio	4,445,773	100	8.3%	Home Visits	Survey Contractor	July 2009
NYS compared to Ohio or NYC			10.1%			
Households in New York City	3,021,588	100	8.3%	Home Visits	Survey Contractor	July 2009
Households in Washington, DC	248,338	100	8.3%	Home Visits	MCA/Impact Evaluation Contractors	July 2009
Households in Houston (optional)	717,945	100	8.3%	Home Visits	MCA/Impact Evaluation Contractors	July 2009
NYC compared to DC or Houston			11.7%			

* Proportion of 50%

VI. Process Evaluation Plan

Historically, NYSERDA's process evaluation activities have used formative approaches to investigate program functioning and identify areas of improvement in a manner very similar to that suggested in the EEPS Evaluation Guidelines, and that will continue as indicated in this plan.

In 2009, when the program is ramping up, NYSERDA plans to conduct a process evaluation of the CFL Expansion Program. The process evaluation will involve in-depth interviews with program staff, implementation contractors, manufacturers, and consumers. In addition, the process evaluation will examine up-stream verification efforts, such as determining the ability of the database to track shipments and sales, and to assess the extent to which incentives are tied to sales rather than shipments. The process evaluation will include interviews with utilities in New York (both IOUs as well as LIPA and NYPA) and DPS staff to understand how the CFL expansion is being coordinated throughout New York. Finally, the process evaluation will involve interviews with program sponsors in other parts of the U.S. to identify best practices.

a. Research objectives

The following are the primary process evaluation objectives. Other objectives may be added as the timing of this research draws closer.

1. To assess how well the program is working from the perspectives of program staff, implementation contractors, consumers, and manufacturers
2. To examine up-stream verification efforts, such as determining the ability of the database to track shipments and sales, and to assess the extent to which incentives are tied to sales rather than shipments
3. To identify coordination issues with New York utilities
4. To determine how the CFL expansion program compares with "best practice" programs in other parts of the U.S.

b. Activities

The activities that will address these process evaluation objectives include surveys and/or in-depth interviews with program staff, implementation contractors, consumers, manufacturers, New York IOUs and DPS staff, and managers of CFL programs in other states. The interviews with program staff and implementation staff will examine how the program operates and what is and is not working well from their perspective. The market surveys of consumers (discussed in the above section on the Market Characterization and Assessment and Impact Evaluation Plan) will include process-related questions about their experiences with CFLs and the program, which will be analyzed for process evaluation purposes. The process evaluation will involve interviews with manufacturers to assess their views of the program, and to ask some market-related questions. The interviews with utilities in New York (both IOUs as well as LIPA and NYPA) and DPS staff will aim to provide an understanding of how the CFL expansion is being coordinated throughout New York. Finally, the process evaluation will involve interviews with CFL program sponsors in other parts of the U.S. to identify best practices.

c. Populations/samples

Table 8 below shows recommended sample sizes. The process evaluation contractor will conduct eight in-depth interviews (IDIs) with program and implementation contractor staff, ten IDIs with participating manufacturers, eight IDIs with New York electric utility representatives and DPS staff, and five IDIs with managers of CFL programs in other parts of the U.S. The process evaluation team will provide survey questions for the New York

consumer interviews that will be conducted for the market characterization and assessment and impact evaluation, and will analyze responses to those questions.

d. Data collection

As shown in

Table 8 below, NYSERDA’s process evaluation contractor will conduct the IDIs (with program and implementation contractor staff, participating manufacturers, New York electric utility representatives and DPS staff, and managers of CFL programs in other parts of the U.S.) during 2009. The RDD survey of households in New York, also conducted for the market assessment and characterization/impact evaluation, took place in December 2008/January 2009 and will be repeated in 2010; since the program had not begun in December of 2008, the process evaluation questions will be included in the early 2010 consumer survey.²³

The in-depth interviews will be conducted by Nexus Market Research staff analysts.

Table 8 – CFL Expansion Program Process Evaluation Survey Specifics

Target Group	Estimated Population Size	Estimated Sample Size	Expected Sampling Precision (90% confidence)*	Data Collection	Survey Administration By	Expected Fielding Dates
Program and implementation staff	≈15	8	NA	IDIs	Process Evaluation Contractor	April-May 2009
Households in NY less NYC & Long Island (see Table 6)	3,118,586	1,000	2.6%	RDD	Survey Contractor	Early 2010)
Households in NYC (see Table 6)	3,021,588	500	3.7%	RDD	Survey Contractor	Early 2010
Participating manufacturers	>34	10	NA	IDIs	Process Evaluation Contractor	April-May 2009
New York electric utilities & DPS staff	≈10	8	NA	IDIs	Process Evaluation Contractor	April-May 2009
Sponsors of CFL programs in other states	≈40	5	NA	IDIs	Process Evaluation Contractor	April-May 2009

* Proportion of 50%

VII. Special Issues

A significant issue with the CFL Expansion Program evaluation is that it involves very intensive data collection and analysis. This will require rapid turnaround of plans and instruments, and flexibility in implementation. Also, as mentioned earlier in this plan, NYSERDA believes much of the data collected through the major market and impact evaluation studies will be of benefit to other NY program administrators and requests that the possibility of joint funding be considered by DPS and the EAG.

²³ As noted previously, a determination regarding the specific timing of the 2010 RDD survey will be made in Fall 2009.

VIII. Schedule and Budget

The budget by year for the CFL Expansion Program evaluation is shown in Table 9; the overall schedule is shown in Table 10. Additional detail has been provided for this evaluation plan budget presentation in order for reviewers to understand the significant investment necessary to conduct the baseline and follow up market and impact evaluation activities outlined herein. Note that the evaluation budget for later in 2010 and 2011 will be determined after the early 2010 evaluation is completed.

Table 9 – CFL Expansion Program Evaluation Budget

Evaluation Element	2009	2010	Total
Market Characterization & Assessment/Impact			
Planning and Management	\$77,420	\$40,258	\$117,678
RDD Survey			
Data Collection	\$90,000	\$93,600	\$183,600
Design and Analysis	\$65,980	\$68,619	\$134,599
Saturation			
Data Collection	\$328,000		\$328,000
Design and Analysis	\$60,726		\$60,726
Retailer Survey			
Data Collection	\$0	\$0	\$0
Design and Analysis	\$0	\$0	\$0
Reporting & modeling	\$79,980	\$62,384	\$142,364
Misc. (Travel for planning, presentations, etc.)	\$5,000	\$5,000	\$10,000
Market Characterization & Assessment/Impact Subtotal	\$707,106	\$269,861	\$976,967
Process Evaluation	\$97,500	\$17,000	\$114,500
Total	\$804,606	\$286,861	\$1,091,467

Table 10 – CFL Expansion Program Schedule

Evaluation Element	2009	2010
Market Characterization & Assessment/Impact		
RDD Survey-NY State, OH, NYC, DC, Houston	X	X
Saturation Survey-NY State, OH, NYC, DC, possibly Houston	X	
Process Evaluation		
Program & Implementation Staff Interviews	X	
Analysis of RDD Survey (fielded by MCAC)		X
IOU & DPS Staff Interviews	X	
Manufacturer Interviews	X	
Best Practices Interviews	X	

IX. NYSERDA Evaluation Process

This evaluation plan is an early, but important step in NYSERDA's evaluation planning and implementation process. It is NYSERDA's understanding that DPS Staff wish to be involved as a reviewer/participant in the following parts of the evaluation process: detailed evaluation plans, project kick-off meetings, workplans, data collection instruments, interim results reports (as applicable), presentation of evaluation results, and draft evaluation reports. NYSERDA will conduct evaluation planning and implementation in an open and transparent manner, and will invite DPS Staff participation in the designated aspects of the process and any others upon DPS' request. Should DPS Staff choose to modify the level or manner of their involvement, NYSERDA should be notified about the change(s). DPS Staff should also choose when and how to involve their evaluation advisor consultant team in NYSERDA's evaluation processes, should directly provide any materials and information necessary for their advisor consultant team to fulfill this role, and should notify NYSERDA about the type and level of advisor consultant involvement.

In cases where a program does not have an existing logic model, or the existing logic model needs to be revisited, logic modeling activities would ideally occur early in the evaluation process after completion and approval of the Detailed Evaluation Plan. NYSERDA's evaluation contractors convene logic model "workshops" with program staff to discuss program inputs, activities, outputs, outcomes, external influences and other elements that need to be documented in the logic model. The evaluation contractors then document these discussions in a brief program theory/logic report, which includes a logic model diagram for the program. NYSERDA will invite DPS Staff to participate in logic model workshops and review draft program theory/logic reports.

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

1. Ongoing communications between the NYSERDA evaluation staff and evaluation contractors to identify issues that need to be brought to the attention of NYSERDA program staff, DPS Staff, and other involved parties.
2. Interim results reports may be generated, sometimes at the request of NYSERDA program staff and sometimes by initiative of NYSERDA's evaluation team and contractors, where early results are required or deemed useful prior to completion of the full evaluation effort.
3. Presentations of draft evaluation results held with NYSERDA evaluation contractors, evaluation team, program staff, and DPS Staff before evaluation reports are written provide feedback on the programs as soon as possible, and provide evaluation contractors with additional perspective and context that will be useful in reporting final recommendations.

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

X. Reporting

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

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

recommendations arising from the evaluation and the status of their consideration or adoption of these recommendations.

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

XI. Total Resource Cost Analysis

Once per year, NYSERDA will update benefit/cost ratios (at a minimum, Total Resource Cost test) for each major program and for the entire portfolio of SBC-funded **New York Energy \$martSM** and EEPS programs. The Total Resource Cost (TRC) test divides the present value of the benefits by the present value of Program and Participant Costs. A benefit-cost ratio greater than 1 indicates benefits exceed NYSERDA and participant costs. The Program Administrator Cost (PAC) test divides the present value of the benefits by the present value of the Program Administrator Costs. A benefit-cost ratio greater than 1 indicates benefits exceed NYSERDA costs. For more detailed definition of benefit/cost terms and a description of NYSERDA's current benefit/cost input sources, including avoided energy, capacity and distribution costs, refer to Appendix A of NYSERDA's September 22, 2008 Energy Efficiency Portfolio Standard Program Administrator Proposal.

The latest evaluated program savings, realization rates, and net-to-gross ratios resulting from the evaluation efforts described in this plan will be used in the annual benefit/cost analysis update.

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