



Detailed Study Plan: Final

National Evaluation of the United States Department of
Energy's State Energy Program

Oak Ridge National Laboratory
Prepared by KEMA Inc and its subcontractors
Fairfax, VA
June 30, 2011

Table of Contents

- 1. INTRODUCTION..... 1**
 - 1.1 PROGRAM DESCRIPTION..... 3
 - 1.1.1 *SEP History*..... 3
 - 1.1.2 *Program Year 2008 v. ARRA Period*..... 5
 - 1.2 EVALUATION OBJECTIVES AND APPROACH..... 8
 - 1.2.1 *Objectives*..... 8
 - 1.2.2 *Overview of Approach*..... 10
 - 1.2.3 *Selected Methodological and Logistical Challenges and Solutions*..... 14
 - 1.3 STRUCTURE OF THE DETAILED STUDY PLAN..... 15
- 2. CHARACTERIZATION OF PROGRAMMATIC ACTIVITIES..... 17**
 - 2.1 ARRA-FUNDED PROGRAMMATIC ACTIVITIES: PY2009 – 2011..... 21
 - 2.1.1 *Sources of Information*..... 21
 - 2.1.2 *Decision Rules for Classifying Programmatic Activities*..... 23
 - 2.2 PY2008 PROGRAMMATIC ACTIVITIES..... 24
 - 2.2.1 *Sources of Information*..... 24
 - 2.2.2 *Decision Rules for Classifying Programmatic Activities*..... 26
 - 2.3 SUB-CATEGORIZATION OF BPACS..... 27
- 3. SAMPLING OF PROGRAMMATIC ACTIVITIES AND EXPANSION OF SAMPLE RESULTS 31**
 - 3.1 OVERVIEW OF SAMPLING APPROACH..... 31
 - 3.1.1 *Total Sample Size*..... 31
 - 3.1.2 *Sampling Frame*..... 31
 - 3.1.3 *Rigor Level*..... 31
 - 3.1.4 *PA Sample Allocation*..... 32
 - 3.2 SAMPLE FRAME..... 34
 - 3.3 SAMPLING TARGETS..... 43
 - 3.4 IMPLEMENTING THE PA SAMPLE DESIGN..... 49
 - 3.4.1 *Misclassification and Multiple Classifications*..... 49
 - 3.4.2 *Independent State-Specific Evaluations*..... 50
 - 3.5 BPAC-SPECIFIC IMPACT CALCULATIONS..... 51
 - 3.5.1 *Portfolio-Level Impact Calculations*..... 52

Table of Contents

3.5.2	<i>Error Bounds</i>	52
4.	ESTIMATION OF ENERGY IMPACTS	54
4.1	INTRODUCTION.....	54
4.1.1	<i>Framework for Specification of Impact Assessment Methods</i>	54
4.1.2	<i>Groupings of Programs for Energy Impact Assessment Planning</i>	55
4.2	EVALUATION PLANS: BUILDING RETROFIT AND EQUIPMENT REPLACEMENT.....	61
4.2.1	<i>Introduction</i>	61
4.2.2	<i>Energy Impacts Assessment Approach</i>	66
4.3	RENEWABLE ENERGY MARKET DEVELOPMENT PROGRAMS.....	78
4.3.1	<i>Introduction</i>	78
4.3.2	<i>Energy Impacts Assessment Approach: Renewable Energy Market Development - Projects</i>	80
4.3.3	<i>Energy Impacts Assessment Approach: Renewable Energy Market Development - Manufacturing</i>	84
4.3.4	<i>Energy Impacts Assessment Approach: Renewable Energy Market Development – Clean Energy Policy Support</i>	88
4.4	INFORMATION AND TRAINING PROGRAMS.....	90
4.4.1	<i>Introduction</i>	90
4.4.2	<i>Energy Impacts Assessment Approach</i>	90
4.5	CODES AND STANDARDS PROGRAMS.....	95
4.5.1	<i>Introduction</i>	95
4.5.2	<i>Estimation of Potential Energy Effects</i>	95
5.	ATTRIBUTION APPROACHES	98
5.1	INTRODUCTION.....	99
5.1.1	<i>Fundamental Research Questions</i>	99
5.1.2	<i>Available Methods for Assessing Attribution</i>	107
5.1.3	<i>Application of Available Methods to Evaluation PAs in Different Groups</i>	112
5.2	ATTRIBUTION APPROACH 1: BUILDING RETROFIT AND EQUIPMENT REPLACEMENT, RENEWABLE ENERGY MARKET DEVELOPMENT – PRODUCTS, INFORMATION AND TRAINING PROGRAMS.....	115
5.2.1	<i>Assessment of Market Actor Response</i>	115
5.2.2	<i>Relative effect of multiple programs</i>	117



Table of Contents

- 5.2.3 *Influence of SEP on Other Programs*.....122
- 5.3 **ATTRIBUTION APPROACH 2: RENEWABLE ENERGY GENERATION AND CAPACITY –
MANUFACTURING, CLEAN ENERGY POLICY SUPPORT, CODES AND STANDARDS**.....123
 - 5.3.1 *Renewable Energy Market Development – Manufacturing*.....123
 - 5.3.2 *Codes and Standards Programs*.....126
- 6. EVALUATION OF CARBON IMPACTS**.....**131**
 - 6.1 ASSESSMENT OF CARBON IMPACTS.....131
 - 6.2 PRESENTATION OF RESULTS.....133
- 7. EVALUATION OF EMPLOYMENT IMPACTS**.....**134**
 - 7.1 ASSESSMENT OF EMPLOYMENT IMPACTS.....134
 - 7.1.1 *Broad Parameters of Jobs Assessment*.....134
 - 7.1.2 *Economic Impact Model for identifying Job Impacts*.....134
 - 7.1.3 *Translating SEP Project Direct Effects into Economic Events*.....138
 - 7.1.4 *Presentation of Job Impacts*.....140
- 8. BENEFIT-COST ANALYSIS**.....**141**
 - 8.1 TYPES OF BENEFIT-COST METRICS.....141
 - 8.2 IMPLEMENTING BENEFIT-COST CALCULATIONS.....143
 - 8.2.1 *SEP RAC Test*.....143
 - 8.2.2 *Net present value of energy savings versus program costs*.....144
 - 8.3 LEVEL OF BENEFIT-COST ASSESSMENT.....144
- 9. TIMELINE**.....**145**



Table of Contents

FIGURES

Figure 1. SEP Funding Allocations by State (PY2008 and ARRA Period).....	6
Figure 2. Key Evaluation Metrics.....	8
Figure 3. Distinguishing Attributes by Refined BPAC.....	19
Figure 4: BPACs and Subcategories.....	29
Figure 5: Percent of PY2008 SEP Budget by BPAC and Subcategory.....	35
Figure 6: Percent of ARRA SEP Budget by BPAC and Subcategory.....	37
Figure 7: Number of Available PAs for Selected BPAC/Subcategory (PY2008).....	39
Figure 8: Number of Available PAs for Selected BPAC/Subcategory (ARRA).....	41
Figure 9: Allocated Sampling Targets by BPAC/Subcategory and Rigor Level (PY2008).....	45
Figure 10: Allocated Sampling Targets by BPAC/Subcategory and Rigor Level (ARRA).....	47
Figure 11. Summary of Gross Energy Savings Estimation and Attribution Approaches by Broad Program Activity Category.....	60
Figure 12. Representation of Energy Savings from Retrofit.....	62
Figure 13. Sample Sizes Required for 90 Percent Confidence Intervals Around Ratio Estimators.....	71
Figure 14. Measurement and Verification for Building Retrofit & Equipment Replacement Group: High Rigor.....	73
Figure 15. Verification for Building Retrofit & Equipment Replacement Group: High Rigor.....	75
Figure 16. Ratepayer Program and SEP Budgets for Selected States (\$ millions).....	103
Figure 17. Applications of Attribution Assessment Methods to Evaluation of PAs in PA Groupings.....	113
Figure 18. Overview of Research on Market Actor Effects.....	117
Figure 19. Considerations in Scoring Market Actor Data on Relative Importance of Multiple Programs.....	121
Figure 20. Forecasts of Measure Sales: Baseline and Actual with Program Support.....	125
Figure 21. Acceleration of Code Adoption.....	127
Figure 22: REMI Economic Forecasting Model – Basic Structure and Linkages.....	136
Figure 23. Identifying Economic Impacts in the REMI Framework.....	137
Figure 24. REEM Framework for Energy Impact Analysis.....	139
Figure 25. SEP National Evaluation Timeline.....	148



Table of Contents

1. Introduction

This document provides a detailed plan for conducting an evaluation of the State Energy Program, a national program operated by the U. S. Department of Energy (DOE). DOE's Office of Weatherization and Intergovernmental Program (OWIP), which manages the State Energy Program, has commissioned this evaluation. Its principal objectives are to develop an independent estimate of key program outcomes:

- Reduction in energy use and expenditures
- Production of energy from renewable sources,
- Reduction in carbon emissions associated with energy production and use, and
- Generation of jobs through the funded activities.

Because of the magnitude and temporal nature of ARRA funding, this evaluation effort has two different but coordinated paths. The contractor team will examine key program outcomes for both the SEP 2008 program year (July 2008 to June 2009) and for ARRA (2009 to present). Based on early feedback from stakeholders and program staff at DOE, this evaluation effort was refocused on 2008 because it will be more likely to characterize the SEP program after the ARRA period, when funding levels return to pre-ARRA levels.

The State Energy Program (SEP) provides grants and technical support to the states and U.S. territories which enables them to carry out a wide variety of cost-shared energy efficiency and renewable energy activities that meet each state's unique energy needs while also addressing national goals such as energy security. Congress created the SEP in 1996 by consolidating the State Energy Conservation Program (SECP) and the Institutional Conservation Program (ICP), which were both established in 1975.

To be counted as part of SEP, an activity must be included in the State Plan submitted to SEP and supported, in part, by SEP funds. While it is not unusual for evaluators to refer to a related set of activities (e.g., multiple energy audits) performed in a single year under a common administrative framework as a "program," such efforts are referred to in this document as "programmatic activities (PA)." Typically, the programmatic activities designed and carried out by the states with SEP support involve a number of actions (e.g., multiple retrofits performed or loans given). In some cases, they combine a number of different types of actions designed to advance the program's objectives, for example: energy audits may be combined with financial incentives such as loans or grants to promote energy efficiency measures in targeted buildings.

In February 2009, the American Reinvestment and Recovery Act (ARRA) was signed into law and allocated \$36.7 billion to the Department of Energy (DOE) to fund a range of energy-related initiatives: energy efficiency, renewable energy, electric grid modernization, carbon capture and storage, transportation efficiency, alternative fuels, environmental management and other energy-related programs. The primary goals for DOE programs funded by ARRA include rapid job creation, job retention, and a reduction in energy use and the associated greenhouse gas emissions; deadlines for fund expenditures were set to ensure that funds were spent within several years. SEP received \$3.1 billion of these funds, which began to be disbursed in late 2009. The deadline for expenditure of all ARRA funds allocated to SEP is April 2012. This program period thus encompasses two and one-half years, spanning SEP's Program Years (PY) 2009 – 2011.¹ By way of contrast, SEP funding in PY2008 was \$33 million.²

Under the American Recovery and Reinvestment Act (ARRA) the amount of funding available to support the states' SEP activities increased dramatically, and as a result, the mix of programmatic activities changed from previous patterns. Once the ARRA funding has been expended, the volume and mix of SEP activities are expected to return to levels typical of the pre-ARRA period. For this evaluation, OWIP has elected to assess the outcomes of programmatic activities for one program year (PY2008) prior to the distribution of ARRA funding, as well as for the full set of programmatic activities that received ARRA support. OWIP believes that this approach will make best use of limited evaluation resources, given that future SEP program years are more likely to resemble pre-ARRA activities than the ARRA-funded activities. These latter will be implemented in Program Years 2009 – 2011. Given the strong differences in volume, scope, and relative priority of policy goals between the pre-ARRA and the ARRA-funded activities, the evaluation team believes it is most appropriate to treat the efforts as separate programs for purposes of sampling state-level activities and estimating national impacts.

The remaining sections of this Introduction provide an overview of SEP as it operated prior to ARRA, the organization and operation of SEP under ARRA funding, the objectives of this evaluation, and its basic methodological approach. Each of these topics is treated in considerable detail in subsequent chapters of this Detailed Study Plan.

¹ In most states, the SEP PY2009 ran from July 1, 2009 to June 30, 2010; PY2010 runs from July 1, 2010 to June 2011; and so on.

² These figures include spending for program administration and emergency energy planning, which are not used for the programmatic activities to be evaluated by this project.

1.1 Program Description

1.1.1 SEP History

Congress created the Department of Energy's State Energy Program in 1996 by consolidating the State Energy Conservation Program (SECP) and the Institutional Conservation Program (ICP). Both programs went into effect in 1975. SECP provided states with funding for energy efficiency and renewable energy projects. ICP provided hospitals and schools with a technical analysis of their buildings and identified the potential savings from proposed energy conservation measures.

Several pieces of legislation form the framework for the State Energy Program:³

- The Energy Policy and Conservation Act of 1975 (P.L. 94-163) established programs to foster energy conservation in federal buildings and major U.S. industries. It also established the State Energy Conservation Program.
- The Energy Conservation and Production Act of 1976
- The Warner Amendment of 1983 (P.L. 95-105) allocated oil overcharge funds—called Petroleum Violation Escrow (PVE) funds—to state energy programs. In 1986, these funds became substantial when the Exxon and Stripper Well settlements added more than \$4 billion into this mix.
- The State Energy Efficiency Programs Improvement Act of 1990 (P.L. 101-440) encouraged states to undertake activities designed to improve efficiency and stimulate investment in and use of alternative energy technologies.
- The Energy Policy Act (EPAcT) of 1992 (P.L. 102-486) allowed DOE funding to be used to finance revolving funds for energy efficiency improvements in state and local government buildings. (However, no funding was provided for this activity.) EPAcT recognized the crucial role states play in regulating energy industries and promoting new energy technologies. EPAcT also expanded the policy development and technology deployment role for the states. Many EPAcT regulations extended through 2000, and we are currently waiting for updates through the National Energy Policy.

³ http://www1.eere.energy.gov/wip/sep_history.html

-
- The American Recovery and Reinvestment Act of 2009 provided \$3.1 billion for SEP formula grants with no matching fund requirements.

1.1.1.1 SEP Goals and Metrics

The State Energy Program (SEP) is a cornerstone of a larger partnership between DOE and the states. SEP program goals therefore reflect the partnership's long-term strategic goals and each energy office's current year objectives.

Goals. The mission of the State Energy Program is to provide leadership to maximize the benefits of energy efficiency and renewable energy through communications and outreach activities, technology deployment, and accessing new partnerships and resources. Working with DOE, state energy offices address long-term national goals to:

- Increase the energy efficiency of the U.S. economy,
- Reduce energy costs,
- Improve the reliability of electricity, fuel, and energy services delivery,
- Develop alternative and renewable energy resources,
- Promote economic growth with improved environmental quality,
- Reduce reliance on imported oil.

The State Energy Program also helps states prepare for natural disasters and improve the security of the energy infrastructure. Specifically, SEP helps states meet federal requirements to:

- Prepare an energy emergency plan,
- Develop individual state energy plans. Each state shares its plan with DOE, sets short-term objectives, and outlines long-term goals.

The State Energy Program outlines this vision and mission in more detail in its Strategic Plan for the 21st Century." ⁴

⁴ http://www1.eere.energy.gov/wip/pdfs/plan_final.pdf

Metrics. Through the State Energy Program, DOE provides a wide variety of financial and technical assistance to the states. States routinely add their own funds and leverage investments from the private sector for energy projects. Some results of the State Energy Program are thus easily measured; for example, energy and cost savings can be quantified according to the types of projects state energy offices administer. Other benefits are less tangible; for example, developing a plan for energy emergencies.

1.1.1.2 Funding Formulas and Competitive Procedures

SEP provides money to each state and territory according to a formula that accounts for population and energy use. In addition to these “Formula Grants,” SEP “Special Project” funds are made available on a competitive basis to carry out specific types of energy efficiency and renewable energy activities (U.S. DOE 2003c). The resources provided by DOE typically are augmented by money and in-kind assistance from a number of sources, including other federal agencies, state and local governments, and the private sector.

1.1.2 Program Year 2008 v. ARRA Period

For program year (PY) 2008, the states’ SEP efforts included several mandatory activities, such as establishing lighting efficiency standards for public buildings, promoting car and vanpools and public transportation, and establishing policies for energy-efficient government procurement practices. The states and territories also engaged in a broad range of optional activities, including holding workshops and training sessions on a variety of topics related to energy efficiency and renewable energy, providing energy audits and building retrofit services, offering technical assistance, supporting loan and grant programs, and encouraging the adoption of alternative energy technologies. The scope and variety of activities undertaken by the various states and territories in PY 2008 was extremely broad, and this reflects the diversity of conditions and needs found across the country and the efforts of participating states and territories to respond to them.

A total of \$33 million in SEP funding was made available during PY2008 to the states and territories as shown in Figure 1. Under the American Recovery and Reinvestment Act (ARRA) the amount of funding available to support the states’ SEP activities increased dramatically and the mix of programmatic activities funded also changed considerably.

Figure 1. SEP Funding Allocations by State (PY2008 and ARRA Period)

State/Territory	PY2008 SEP Formula Grant Allocation	SEP ARRA Obligations
Alabama	\$517,000	\$55,570,000
Alaska	\$250,000	\$28,232,000
American Samoa	\$160,000	\$18,550,000
Arizona	\$476,000	\$55,447,000
Arkansas	\$403,000	\$39,416,000
California	\$2,151,000	\$226,093,000
Colorado	\$518,000	\$49,222,000
Connecticut	\$493,000	\$38,542,000
Delaware	\$223,000	\$24,231,000
District of Columbia	\$212,000	\$22,022,000
Florida	\$1,135,000	\$126,089,000
Georgia	\$734,000	\$82,495,000
Guam	\$167,000	\$19,098,000
Hawaii	\$233,000	\$25,930,000
Idaho	\$259,000	\$28,572,000
Illinois	\$1,398,000	\$101,321,000
Indiana	\$800,000	\$68,621,000
Iowa	\$472,000	\$40,546,000
Kansas	\$422,000	\$38,284,000
Kentucky	\$539,000	\$52,533,000
Louisiana	\$620,000	\$71,694,000
Maine	\$298,000	\$27,305,000
Maryland	\$615,000	\$51,772,000
Massachusetts	\$753,000	\$54,911,000
Michigan	\$1,177,000	\$82,035,000
Minnesota	\$716,000	\$54,172,000
Mississippi	\$378,000	\$40,418,000
Missouri	\$656,000	\$57,393,000
Montana	\$244,000	\$25,855,000
Nebraska	\$321,000	\$30,910,000
Nevada	\$279,000	\$34,714,000
New Hampshire	\$280,000	\$25,827,000
New Jersey	\$964,000	\$73,643,000
New Mexico	\$297,000	\$31,821,000

State/Territory	PY2008 SEP Formula Grant Allocation	SEP ARRA Obligations
New York	\$1,941,000	\$123,110,000
North Carolina	\$750,000	\$75,989,000
North Dakota	\$232,000	\$24,585,000
Northern Marianas	\$160,000	\$18,651,000
Ohio	\$1,311,000	\$96,083,000
Oklahoma	\$463,000	\$46,704,000
Oregon	\$427,000	\$42,182,000
Pennsylvania	\$1,336,000	\$99,684,000
Puerto Rico	\$412,000	\$37,086,000
Rhode Island	\$258,000	\$23,960,000
South Carolina	\$463,000	\$50,550,000
South Dakota	\$226,000	\$23,709,000
Tennessee	\$628,000	\$62,482,000
Texas	\$1,858,000	\$218,782,000
Utah	\$327,000	\$35,362,000
Vermont	\$226,000	\$21,999,000
Virgin Islands	\$174,000	\$20,678,000
Virginia	\$742,000	\$70,001,000
Washington	\$585,000	\$60,944,000
West Virginia	\$366,000	\$32,746,000
Wisconsin	\$740,000	\$55,488,000
Wyoming	\$215,000	\$24,941,000
Total	\$33,000,000	\$3,069,000,000

1.2 Evaluation Objectives and Approach

1.2.1 Objectives

The overall objective of this evaluation is to develop independent, quantitative estimates of key program outcomes for the largest programmatic activities accounting for at least 80 percent of funding for each period of study, and aggregated to selected groups of Broad Programmatic

Activity categories (BPAC) that share common energy savings mechanisms as described in Section 4.

Figure 2 lists the key metrics to be estimated along with elements of the working definitions DOE has assigned to them for purposes of this evaluation.

Figure 2. Key Evaluation Metrics

METRIC CATEGORY/Metric	Elements of the Working Definition
ENERGY SAVINGS Annual energy savings	<ul style="list-style-type: none"> • Fuel units such as kWh/Year for electricity, therms/Year for natural gas, gallons of oil for all energy resources typically procured from a commercial supplier • Percentage of pre-program energy use • Weather normalized, that is, adjusted as needed to local weather conditions in a typical meteorological year (TMY)
Lifetime energy savings	<ul style="list-style-type: none"> • Annual savings realized over the effective useful life (EUL) of the measures installed, that is: the period of time over which savings are expected to be achieved • See Section 4 for a description of the proposed implementation of the EUL
Electric demand savings	<ul style="list-style-type: none"> • Effect of measures evaluated on local electric system peak demand • May be estimated using one or more readily available approaches, such as application of coincidence factors or load shapes for the measures in question

Figure 2 (continued). Key Evaluation Metrics

METRIC CATEGORY/Metric	Elements of the Working Definition
RENEWABLE ENERGY CAPACITY AND GENERATION Capacity	<ul style="list-style-type: none"> Installed capacity of renewable energy facilities developed with the assistance of or otherwise facilitated by SEP programmatic activities Measured as kW installed for photovoltaic, wind, small hydroelectric, tidal energy, and bio-fuel powered generating facilities; BTU/hour for solar hot water and bio-fuel thermal facilities
Annual Renewable Energy Generation	<ul style="list-style-type: none"> Annual energy supplied by renewable energy facilities developed with the assistance of or otherwise facilitated by SEP programmatic activities, as denominated in units of the fuel displaced or coal equivalent if not displaced
Lifetime Renewable Energy Generation	<ul style="list-style-type: none"> Amount of energy supplied by renewable energy facilities developed with the assistance or otherwise facilitated by SEP programmatic activities over the effective useful life of the facilities
ENERGY COST SAVINGS Annual energy cost savings	<ul style="list-style-type: none"> Value of annual energy savings, demand reductions, and annual renewable energy generation at current customer costs
Lifetime energy cost savings	<ul style="list-style-type: none"> Customer value of annual energy savings and demand reductions at current customer costs over effective useful life
CARBON EMISSIONS REDUCTIONS Annual CO2 Emissions Avoided	<ul style="list-style-type: none"> Tons per year of avoided CO2 emissions resulting from: (a) reduced use of fossil fuels due to program activities (i.e., reduced direct use of natural gas and fuel oil, and reduced use of electricity generated from fossil fuels), and (b) reduced use of fossil fuels due to replacement of fossil fuel-generated electricity with electricity generated from renewable sources
Lifetime CO2 Emissions Avoided	<ul style="list-style-type: none"> The sum of annual CO2 emissions avoided as defined above over the useful lives of the measures evaluated
DIRECT AND INDIRECT JOB IMPACTS Direct Job Impacts (Created or Retained)	<ul style="list-style-type: none"> National-level employment activity caused by spending on SEP/ARRA staff and implementation teams to implement SEP funded projects (involved in administration, on-site audits/installation, trainings) to be stated in full-time-equivalents (FTEs) for annual average impact or job-years
Indirect Job Impacts (Created or Retained)	<ul style="list-style-type: none"> National-level employment activity caused by increased market movements in the areas impacted (contractors, retail, wholesale, transportation, etc.) to be stated in full-time-equivalents (FTEs) for annual average impact or job-years National-level employment activity related to longer term jobs that are the results of the spending of the energy savings in the economy into the future to be stated in full-time equivalents (FTEs) for annual average impact or job-years

DOE expects that the analysis conducted to quantify the evaluation metrics listed in Figure 2 will help to identify lessons learned that can be applied to improve the outcomes and cost effectiveness of future SEP operations.

1.2.2 Overview of Approach

The basic steps or stages in the evaluation will be as follows.

- **Characterize the full set of PY2008 and ARRA-funded programmatic activities in terms of Broad Program Activity Categories (BPACs) and measures of size.** In terms of the evaluation, the principal objectives of this step are to:
 - Develop the sample frame from which the individual PAs to be evaluated will be selected, and based on which results for individual PAs will be expanded to the full program.
 - Provide input data to support sample design, including the definition of subcategories in addition to Program Year and BPAC grouping and the allocation of sample resources to final set of sample subcategories.
 - Develop the information needed to expand the results from the sampled PAs to estimate total impacts for the BPAC Groups, PY 2008 Programmatic Activities, and ARRA-funded programmatic activities.
 - Gather information on the level and quality of available program documentation, which will be used to make final determinations of evaluation approaches to be taken in regard to specific BPACs.
- **Develop the sample of individual PAs for evaluation.** The KEMA team will select a sample of at least 82 individual PAs from more than 450 in operation during PY2008 and 575 ARRA-funded PAs. See Section 3 for a description of the objectives, methods, and preliminary design of the PA sample selection process. Once a PA has been selected into the sample, the KEMA team will deploy the evaluation in the following steps.
- **Assess evaluability of the sampled individual PAs.** The Evaluation Team will need some specific pieces of information in order to determine whether a PA that has been selected into the sample can be evaluated at the assigned level of rigor. These are as follows.

-
- a. *Match of actual program operations to the BPAC definition.* As discussed below, the KEMA team has developed detailed working definitions for each BPAC. If, upon selection and detailed review of activities, we find that a PA has been misclassified, it will be evaluated consistent with its actual activity. Its expansion weight—or factor used to project an estimate to the population—will be based on the BPAC it was selected from.
 - b. *Progress in implementation.* In order to carry out high- or medium-high-rigor evaluations, the program needs to have resulted in a sufficient number of the targeted actions, such as completion of retrofit projects or installations of renewable energy equipment, for a sample to be drawn and tested by December 2011.
 - c. *Quality and availability of program records.* For high- and medium-high-rigor evaluations, it will be necessary to contact participants in the program. In most cases we will need to be able to characterize the services that participants received from the program at the individual level. If such records are not available at the time of PA selection and cannot, in the evaluator's judgment, be reconstructed within schedule and budget constraints, then the PA will be dropped from the sample and a substitute selected. If a large proportion of the PAs in a BPAC have insufficient data to support a medium high rigor evaluation, it may be necessary to reduce the rigor level for the BPAC.

In the evaluation plans below we identify the criteria we will use to assess evaluability for each BPAC.

- **Prepare evaluation plans for the BPACs of selected individual PAs.** Once the evaluability of the selected PA has been established, the next step will be to incorporate that PA into its associated BPAC plan that takes into account its specific goals and objectives, market environment, activities and service offerings, and the quality of its tracking records as necessary. As described in Section 4 below, the individual BPAC evaluation plans will be short, highly structured documents that specify the type and amount of data collection to be carried out, the types of analytic approaches to be applied, the staff and subcontractors to be used, the labor and direct costs required, and the implementation schedule. These plans are meant to serve primarily as a tool for managing overall project resources and for quality control.
- **Estimate the energy impacts of the selected individual PAs.** For each selected individual PA, the KEMA team will carry out an assessment of energy impacts. That is,

we will quantify the energy savings, renewable energy capacity and generation, and energy cost savings metrics listed in Figure 2 at the level of rigor specified by DOE. For this evaluation, DOE has identified three levels of rigor for assessment of energy impacts:

- High-rigor evaluations require verification of savings through best practice methods, particularly methods recognized in the *California Evaluation Protocols*, DOE's *Impact Evaluation Framework for Technology Deployment Programs*, and the *International Performance Measurement and Verification Protocol*. These methods include on-site verification and/or performance monitoring of a sample number of projects supported by the program, whole building utility meter billing analysis, surveys of participants and nonparticipants, and combinations of building simulation modeling and other engineering analysis with the first two methods.⁵ In some cases these verification methods may be mixed with less intensive approaches such as file review and telephone contact with program participants to increase sample size. Sample results are expanded to the population using statistical methods, such as ratio estimation or regression analysis.
- Medium-high-rigor evaluations require verification of savings with individual participants, using less intensive data collection and analysis methods than those prescribed for high rigor. All input data may be collected through telephone contact with participants, supplemented by review of program documentation. These data are then combined with documented input assumptions and applied to standard engineering formulae to estimate savings for all or a sample of participants.⁶ On-site data collection, if used at all in medium rigor evaluations, will be applied either in exceptional cases, such as when a single project represents a large portion of potential savings for the PA, or where needed to support key assumptions used in the engineering-based assessments. Sample sizes will also be smaller in the medium-high-rigor assessments.
- Medium-low-rigor evaluations will not include any primary data collection from individual program participants to estimate savings. Rather it will combine

⁵ Given that the majority of sampled PAs will be municipal, commercial, and industrial end users for whom billing analysis is not particularly accurate, and that electric utilities are not sponsoring this study, we anticipate that it will not be possible to collect billing data from a sufficient number of participating sites to support billing analysis.

⁶ These approaches are commonly referred to as engineering-based assessment or statistically-adjusted engineering assessment.

information that can be gained from program records with secondary sources and engineering-based methods to generate energy savings estimates.

See Section 4 for details of high, medium-high, and medium-low rigor energy impact assessment approaches to be applied in regard to specific BPAC groups.

- **Assess the attribution of estimated energy impacts to the individual PAs.** For each selected individual PA, the KEMA team will carry out an analysis to assess the portion of estimated energy impacts that were attributable to the SEP programmatic activities under review, as opposed to other influences such as general developments in the market or the activities of other organizations offering similar kinds of programs or services. For assessing the attribution effects of the SEP, because multiple funding sources are common, impacts must be attributed to SEP and other sources. The ramifications of this are as follows:
 - Attribution of effects must be assessed separately for each individual programmatic activity study.
 - A multi-step attribution approach will be used to include logic models, model validation, cause and effect relationships, funding stream analysis, behavior change assessment, and other established techniques to quantify effects.
 - An examination of what SEP caused to happen will need to account for program-induced capacity developed over time.

See Section 5 for a discussion of our general approach to attribution and its application to evaluation of PAs in specific BPAC Groups.

- **Estimate effects of individual PAs on carbon emissions.** The contractor team will use estimates of annual and lifetime energy savings attributable to the program as inputs to a model that estimates carbon emissions reductions based on the carbon content of fossil fuels and electricity consumption avoided. See Section 6 for a description of this analysis.
- **Estimate effects of individual PAs on employment.** The energy savings estimates will be combined with other program information, such as matching funds contributed, participant expenditures for labor and materials, and direct program expenditure as inputs into a regional economic model to estimate employment impacts. See Section 7 for a description of these analyses.

- **Estimate costs and benefits.** Program reporting guidelines⁷ require that sponsors use only one cost-effectiveness test, designated the SEP Recovery Act Cost (SEP RAC) Test which is computed as source BTUs saved per \$1,000 in program expenditure or investment. The SEP Recovery Act Financial Assistance Funding Opportunity Announcement specified that states should seek to achieve annual energy savings of 10 million source BTUs per \$1,000 of program investments. See Section 8 for further detail on benefit-cost analysis.

Once the individual PA evaluations have been completed and reviewed by the senior contractor team for accuracy and completeness, the effort will shift to aggregation of sample results to the national level and interpretation of findings. KEMA and its subcontractors will expand the sample results to the most well-funded BPACs, using the relationship between verified metrics for the sample PAs and information on measures of size (funding).

1.2.3 Selected Methodological and Logistical Challenges and Solutions

The remainder of this Detailed Study Plan contains separate and fairly extensive chapters on each of the key sets methodological requirements: assessment of energy impacts, program attribution, carbon emission reductions, job creation, and general benefits and costs. Our proposed methods take into account a number of considerations in addition to the challenges posed by the scope of SEP activities and DOE's evaluation objectives. Principal among these are the following.

- **Paperwork Reduction Act Requirements.** The Paperwork Reduction Act (PRA) requires that all data collection instruments and protocols that will be administered to 10 or more "people of the general public, including federal contractors" be reviewed and approved by the Office of Management and Budget (OMB). The 11-step process includes three periods of public comment totaling 120 days, and OMB has 60 days to make its final approval decision following the close of public comments. The process will likely require six to nine months to complete. Once data collection instruments have been reviewed and approved by OMB, they may not be changed except within prescribed bounds to facilitate their administration in a variety of settings. Given these constraints, the KEMA team has designed our overall research effort to optimize the number of data collection forms and protocols that will require OMB review.
- **Evaluations of ARRA-funded SEP programs funded and sponsored by individual states.** As of the submission of this Detailed Study Plan, KEMA is aware of a number of

⁷ Department of Energy, SEP Program Notice 10-07, Attachment 3.

evaluations of ARRA-funded SEP programmatic activities being conducted by state energy offices and other program sponsors.⁸ The KEMA team will coordinate with these efforts to avoid sampling programmatic activities that are being evaluated by the states. The KEMA team will assess the methods being used by the states to determine whether they meet the rigor levels and employ methods approved by DOE. In those cases, we will incorporate the results of these studies into the estimate of national impacts, using the sample stratification and weighting system described in Section 3. Any recommendations for importation of results from other efforts will be submitted to DOE for review and approval prior to implementation.

1.3 Structure of the Detailed Study Plan

The remainder of the Detailed Study Plan is structured into the following sections.

- **Section 2: Characterization of Programmatic Activities.** This section presents the methods and results of the evaluation team's efforts to classify and characterize SEP programmatic activities at the state level for PY2008 and under ARRA funding. The results of this analysis form the basis of our proposed sampling plan.
- **Section 3: Sampling Plan and Expansion of Sample Results.** This section presents the approach for selecting the programmatic activities to be evaluated, including sample segmentation, allocation of sample to segments defined by BPAC and rigor level, estimation of expected sampling error at the BPAC and program levels, and sample selection procedures. This section also summarizes methods to expand the findings from the sample programmatic activities to the full set of programmatic activities.
- **Section 4: Estimation of Energy Impacts.** This section provides a summary of the methods to be applied in estimating energy savings and renewable energy generation associated with each Broad Program Activity Category.
- **Section 5: Attribution Assessment.** This section presents the basic strategies and methods that will be applied to assess the attribution of observed outcomes to the

⁸ As of the date of this study plan, the list of states is still growing and the scope and rigor levels are highly variable; however, KEMA has received indication from various sources that the following states are engaging in some form of ARRA evaluation activity: California, New York, Missouri, Maine, Pennsylvania, Ohio, Wisconsin, Washington State, Oklahoma, Delaware, New Hampshire, Utah, Nevada, Massachusetts, and Georgia.

effects of the sample programmatic activities, and their application to specific kinds of PAs in the BPACs.

- **Section 6: Evaluation of Carbon Impacts.** This section presents the methods that will be applied to quantify national-level carbon reductions for the BPACs evaluated and for the program as a whole.
- **Section 7: Evaluation of Employment Impacts.** This section presents the methods that will be applied to quantify net jobs created for the BPACs evaluated and for the program as a whole.
- **Section 8: Benefit-Cost Analysis.** This section presents the methods that will be used to collect and analyze benefit and cost information at both the individual PA and aggregated levels. It covers the full range of benefit and cost metrics, as well as the relevant benefit-cost test.
- **Section 9: Project Schedule.** This section presents the schedule of project activities and deliverables.

2. Characterization of Programmatic Activities

Transforming the PY2008 and ARRA program data into a format that can support evaluation research is a key task in the design and execution of this study plan. The program data will serve as the backbone for all major evaluation study elements, including the following:

- An evaluability assessment
- Sample development and stratification
- Sample expansion of findings to the population
- Methodological development for gross and net savings estimation

The KEMA team received the program tracking data from DOE for PY2008 (maintained in the WinSaga database system) and from ARRA (in the PAGE database system) and conducted an extensive review. On balance, neither dataset was well suited to support the kind of analyses required under this evaluation effort. The PAGE database for ARRA is uniformly more complete and internally consistent than is the WinSaga database for PY2008. However, the data content of each database lacked an organizational structure for the first key task of the evaluation team: sorting and classifying the programmatic activities into categories established by past SEP evaluation efforts and according to the requirements of the Statement of Work. KEMA manipulated the data into a structure that was organized by programmatic activity and produced data management tools that facilitated supplemental data collection for this project.

The Statement of Work for this study provides the following guidance for BPACs based on past SEP evaluation research and the metric categories provided in the Funding Opportunity Announcement (FOA) for the SEP grants under ARRA.⁹ Those original sixteen BPACs specified in the Statement of Work are as follows:

- Retrofits
- Renewable energy market development
- Loans, grants, and incentives
- Workshops, training, and education
- Building codes and standards
- Industrial retrofit support
- Clean energy policy support

⁹ Funding Opportunity Announcement Number DE-FOA-0000052 issued by National Energy Technology Laboratory, State Energy Program Grants (Issue Date: April 24, 2009), Pages 39 to 40 (Section 10.2A).

-
- Traffic signals and controls
 - Carpools and vanpools
 - Technical assistance to building owners
 - Commercial, industrial, and agricultural audits
 - Residential energy audits
 - Government and institutional procurement
 - Energy efficiency rating and labeling
 - Tax incentives and credits
 - New construction and design

As the first key step, KEMA team members worked collaboratively with key study authors of past SEP evaluation research¹⁰ to develop standards and decision rules for the sorting and classification tasks. Since many of the activity descriptions provided in the SOW are derived from the FOA, contractor staff reviewed the FOA to ensure that the standards used to classify the programmatic activities were consistent with the FOA's intent. KEMA then established a set of distinguishing attributes for the BPACs based on the information obtained from SEP researchers and the FOA language to ensure consistency in assignment across the team. In some cases, we thought it was necessary to decompose some of the BPACs by market segment or program delivery mechanism. Additionally, DOE directed the contractor team to bundle PAs relating to the Workshops, Education and Training (WET) BPAC into the remaining BPACs, removing the WET-related PAs as a BPAC altogether.

Finally, we began the process of assigning programmatic activities to the BPACs and the process thereafter was somewhat iterative. As the KEMA team learned more about the actual programmatic activities, the distinguishing attributes of the BPACs were further refined and the classifications were recast.

¹⁰ Specifically, KEMA received guidance from Martin Schweitzer of ORNL and Nick Hall of TecMarket Works.

summarizes the distinguishing attributes of each BPAC as they have been settled through the iterative process.

Figure 3. Distinguishing Attributes by Refined BPAC

BPAC	Distinguishing PA Attributes Relevant to Primary BPAC Designation
<p>Building Retrofits</p>	<ul style="list-style-type: none"> • Provides financial incentives for building retrofit and equipment replacement projects in non-residential and residential buildings. • Non-residential projects typically identify specific facilities, or facility owners in the grant application or PA description. • Residential programs do not identify specific projects, facilities, or customers.
<p>Technical Assistance</p>	<ul style="list-style-type: none"> • Provides technical assistance other than audits for building retrofit or equipment replacement projects: e.g. technical studies for specific improvements, building modeling, project financial analysis, support in negotiating with contractors. • Open to commercial, industrial, and agricultural facility owners or specified subgroups thereof. • May be combined with financial incentives.
<p>Energy Audits: Commercial, Industrial and Agricultural</p>	<ul style="list-style-type: none"> • Provides funding for or direct services for energy audits of commercial, industrial, and agricultural facilities. Could range from simple checklist to investment-grade audits, mostly involves onsite delivery. • Audits are oriented to identifying cost-effective building retrofit and/or equipment replacement projects. • May be combined with financial incentives.
<p>Energy Audits: Residential</p>	<ul style="list-style-type: none"> • Provides funding for or direct services for energy audits of residential facilities. Could range from on-line to on-site audits. • Audits are oriented to identifying cost-effective building retrofit and/or equipment replacement projects. • May be combined with financial incentives.
<p>Renewable Energy Market Development</p>	<ul style="list-style-type: none"> • Provides financial incentives and/or technical assistance to support the development of renewable energy facilities including: solar, wind, biomass, small hydro. • Includes PAs that develop or expand existing manufacturing capacity for renewable energy equipment or components. • At least some portion of the output of the new or expanded capacity is intended for domestic installation.
<p>Clean Energy Policy Support</p>	<ul style="list-style-type: none"> • Develops and obtains legislative, executive, or regulatory approval for policies to facilitate the completion of renewable energy facilities. Examples might include statewide zoning laws, feed-in tariffs, favorable back-up tariffs, renewable portfolio standards.
<p>Transportation</p>	<ul style="list-style-type: none"> • Provides training, financial support, technical assistance, marketing assistance, and/or administrative assistance to facilitate the development and operation of car and van pools.

BPAC	Distinguishing PA Attributes Relevant to Primary BPAC Designation
	<ul style="list-style-type: none"> • Supports capital improvements to support substitution of renewable fuels or electricity for conventional transportation fuels. • Supports improvements to fleet vehicle efficiency and operations. • Includes traffic signal optimization and control upgrades that reduce idling times.
<p>Traffic Signals</p>	<ul style="list-style-type: none"> • Only provides incentives and technical support for LED traffic signals retrofit and replacement. • Controls upgrades that aim primarily at reducing idling times are included in the expanded Car Pool and Van Pool BPAC – now Transportation.
<p>Building Codes and Standards</p>	<ul style="list-style-type: none"> • Provides marketing support for products that meet the higher energy efficiency standards. • Provides training to vendors in marketing and installation of products that meet the higher energy efficiency standards. • Provides technical and administrative support for the development of more energy-efficient state and federal equipment standards and building codes. • Provides training and technical services to strengthen enforcement of the energy elements of state building codes.
<p>Energy Efficiency Rating and Labeling</p>	<ul style="list-style-type: none"> • Provides technical and administrative support for the development of energy efficiency ratings of energy-using equipment or buildings. • Provides marketing services to build customer awareness of the subject energy efficiency ratings. • Provides training and technical services to build vendor awareness and use of energy efficiency ratings in their business activities.
<p>Government, School and Institutional Procurement</p>	<ul style="list-style-type: none"> • Provides technical and administrative support for government initiatives to purchase energy-efficient equipment or energy-efficient design services.
<p>New Construction and Design</p>	<ul style="list-style-type: none"> • Provides technical and administrative support for the development of energy efficiency ratings of energy-using equipment or buildings. • Provides marketing services to build customer awareness of the subject energy efficiency ratings. • Provides training and technical services to build vendor awareness and use of energy efficiency ratings in their business activities.
<p>Loans, Grants, and Incentives</p>	<ul style="list-style-type: none"> • Provides financial incentives for building retrofit and equipment replacement projects in non-residential buildings. • Does not identify specific projects, facilities, or customers. • Incentives allocated according to an open application process for eligible customer groups. • Financial incentives are the principal program offering, but may be combined with others such as audits.
<p>Tax Incentives and Credits</p>	<ul style="list-style-type: none"> • Provides or facilitates access to state and federal tax credits for building retrofit or energy-efficient equipment replacement projects in residential facilities. • May be combined with technical services.

BPAC	Distinguishing PA Attributes Relevant to Primary BPAC Designation
Allocations To Be Removed from Sample Frame	
Administration	<ul style="list-style-type: none"> • General administration and back-office support for market title activities.
Energy Emergency Planning	<ul style="list-style-type: none"> • All activities related to mitigating energy disruptions during emergency situations. • Includes monitoring energy supplies, demand, and prices and communicating this information to the public.

2.1 ARRA-funded Programmatic Activities: PY2009 – 2011

The specific BPAC data classification activities for ARRA are described first for several reasons. DOE provided these data first, exported from the PAGE database.¹¹ Additionally, these data are more complete and consistent across the states.

This section describes KEMA's approach to developing the frame for analysis for the ARRA period. First, we describe the sources of information, the decision rules, and then some basic descriptive statistics on the results of those classification activities. Data quality issues are addressed throughout this section at each step in the process.

2.1.1 Sources of Information

DOE delivered the PAGE database complete through the third quarter of 2010 in the form of five separate Excel spreadsheets. KEMA analyzed the data structure and established the relationships between each of the spreadsheets and imported the data into Microsoft Access. KEMA also interviewed key DOE staff on the key data contents and reviewed them for any value in classifying programmatic activities according to the BPACs. To complete the BPAC sorting and classification task, the data required the following:

- A unique list of Programmatic Activities: The third quarter (Q3) 2010 PAGE data contained 443 Market Titles. Upon review of the data, these could be derived from either the Market Title data or the Activity data which, in turn, represented a complete set of finer composite data for each Market Title parent.

¹¹ PAGE stands for "Performance and Accountability for Grants in Energy."

-
- Funding data associated with each Programmatic Activity: The funding allocation could either be reported at the Market Title level or as the sum of funding for all Activities within a given Market Title.
 - Descriptive information to assist in the classification process: KEMA determined that the data field most closely aligned with the original 16 BPACs identified in the SOW is called the “Main Metric Area” and was associated with the Market Title records, but not the lower level Activity records.

KEMA performed a join query operation to associate all Activity level data with the associated Market Title data at the parent level.

The subsequent sorting and classification process that followed had four steps:

1. Preliminary BPAC data match: KEMA developed a matching algorithm to assign a preliminary BPAC to each Activity based on the Main Metric Area based on how closely it aligned with a given BPAC. In most cases the match was exact. For BPACs having no reasonable match in the Main Metric Area data, no preliminary BPAC was assigned.
2. PAGE Activity Record Review: KEMA regional coordinators reviewed the detailed record data for each activity and either confirmed or reassigned the preliminary BPACs as appropriate.
3. Internet Research on Programmatic Activities: KEMA regional coordinators organized teams of analysts to perform internet research on various programmatic activities and supplement information to the PAGE data as appropriate. KEMA updated preliminary BPAC assignments based on any new information uncovered.
4. Interviews with the assigned state DOE Project Officers: To minimize burden on, and build a rapport with, the DOE Project Officers (POs), the KEMA regional coordinators were assigned states to ensure that DOE Project Officers were communicating with only one KEMA staff member—namely the regional coordinator. KEMA developed a brief interview guide in consultation with ORNL/DOE and reviewed all programmatic activities and the BPAC assignments with the DOE POs.
5. Verification of PA data by State Energy Program staff. Because the sample data was developed from the Q3 2010 PAGE data, the KEMA contractor team verified the status

of all PAs with the States' Energy Offices, such as whether funding amounts changed, the PA was dropped, and in some cases whether the BPAC assignment had changed.

Throughout this iterative process described above, regular meetings informed the process to ensure consistency in the BPAC assignments across regional coordinators, and to refine and tighten the distinguishing attributes for each BPAC.

2.1.2 Decision Rules for Classifying Programmatic Activities

While the decision rules are presented above, KEMA needed to maintain some basic principles in its BPAC assignments because many, if not most, programmatic activities have elements of multiple BPACs in them. The basic principles were:

- Assign the BPAC that most fits the programmatic activity.
- Assign the highest level rigor possible that reasonably fits the programmatic activity.
- Assign a secondary or tertiary BPAC if a programmatic activity exhibits such strong supporting elements.
- Assign "Administration" as a BPAC for funded activities that are primarily administrative in nature and have no programmatic feature that would deliver energy savings.

As a result of the iterative process and basic principles described above, KEMA made some key refinements to the BPAC distinguishing attributes, including the following:

- If the programmatic activity included a condition for a retrofit component, it was assigned to a retrofit BPAC. For example, if an audit program only provided funding for the audit if the retrofit was performed, this would be assigned as a Building Retrofit.
- Because the gross estimation procedure and sampling unit for LED traffic signal upgrades—an efficiency measure—is so different from traffic control systems designed to reduce idling times and emissions, programmatic activities that primarily focused on reducing transportation emissions were assigned to the Transportation category, expanded from Carpools and Vanpools.

2.2 PY2008 Programmatic Activities

In this section, we describe how KEMA developed the frame for the PY2008 programmatic activities. Following the organization above, first we present the sources of information, the decision rules, and then some basic descriptive statistics on the results of those classification activities. Data quality issues are addressed throughout this section at each step in the process.

2.2.1 Sources of Information

Based on the knowledge gained by KEMA through the PAGE database development process, KEMA supplied DOE with general specifications for the requested PY2008 data. DOE was required to complete a fairly extensive manual process to extract and build the requested program data sets from the WinSaga Program Tracking Database. DOE delivered the data sets covering the PY2008 period in six separate Excel spreadsheets. KEMA analyzed the data structure and established few relationships between each of the data sets, but determined that two data sets, namely the Market Title data set and the Metrics data set, contained the information required. KEMA also interviewed key DOE staff on the key data contents and reviewed them for any value in classifying programmatic activities according to the BPACS. Like the PAGE data, to complete the BPAC sorting and classification task, the data required the following:

- A unique list of Programmatic Activities: Upon review of the data, these could be derived from the Market Title data which did not have any associated data for further disaggregation. The PY2008 WinSaga data contained 578 unique Market Titles covering 55 states/territories. KEMA had no way of verifying that the PY2008 records of Market Titles are complete and one state in particular, Maryland, had no programmatic activity data associated with it whatsoever.
- Funding data associated with each Programmatic Activity: KEMA reviewed the data with a DOE program manager who explained several critical details which have direct impacts on the evaluability of the PY2008 program as well as the sample planning. Funding data from WinSaga differs from the PAGE data in the following ways:
 - o Data pulled from any given program year do not add to the amount allocated to that program year. For PY2008, states were allocated \$33 million for the SEP program; however, data on funding exceeded \$62 million.

- o Market Titles frequently have no data associated with them. Since states have five years to use the SEP funding, unspent funding can roll over to the next program year, or be allocated over several years. KEMA observes that nearly one-third (35 percent) of 579 Market Titles had no funding data associated with them.
- o States' planning cycles can differ substantially from the July 2008 to June 2009 SEP Program Year, which could impact how the funding by Market Title was maintained in the WinSaga database. For example, states were reporting their individual fiscal year, calendar year, or program year cycles which was often different than the federal program year cycle.
- Descriptive information to assist in the classification process: KEMA determined that the data field most closely aligned with the original 16 BPACs identified in the SOW is called the "Metric Area," but this was not associated with the Market Title records. In the exported WinSaga data, the Market Title and Metric Area data are completely unrelated; however, KEMA was able to determine that a one-to-many relationship generally existed between Market Titles (unique) and Metric Areas (unique to a Market Title).

To render the data usable, KEMA manually built a data set using the following steps:

1. Establish the unique set of Market Titles by comparing with the Metric data. The Metric data has 1,642 records covering 46 states/territories with market title data, but the number of records exceeded the number of unique market titles by almost three to one. The market title data matching process was only partially successful, requiring KEMA to manually match a large proportion of the data sets by market title. In the end, KEMA only discovered one market title in the Metrics data set that could not be matched to the Market Title data set and added in that particular record (making 579 records in the Market Title database).
2. Preliminary BPAC data match: KEMA developed a matching algorithm to assign a preliminary BPAC to each Metric Area in the Metrics data set based on how closely it aligned with a given BPAC. In most cases the match was exact. For BPACs having no reasonable match in the Metric Area data, no preliminary BPAC was assigned. In many cases, a given market title had multiple matched BPACs.
3. Narrowing of Metric data: Although the Metrics data are incomplete relative to the unique records in the Market Title data set, for the matched data records, a one-to-many

relationship existed between the Market Title data and the Metrics data set. KEMA manually selected a unique record in the Metric data set to be merged into the Market Title dataset based on the matched BPAC with the highest incidence per market title.

4. Merge the Metrics and Market Title data sets: KEMA merged the narrowed Metrics data set into the Market Title dataset, creating a data frame with descriptive information on programmatic activity and the associated metric.

KEMA staff reviewed the detailed merged record data for each activity and either confirmed or reassigned the BPACs as appropriate, using the standards and distinguishing attributes established in developing the ARRA data from above. The review process had several iterations and included several KEMA staff involved in the BPAC assignments to ensure consistency in BPAC assignment between program periods.

Like the ARRA data classification process, the contractor team verified the data accuracy through interviews with NASEO Regional Coordinators and the assigned state DOE Project Officers. KEMA first reached out to the NASEO Regional Coordinators—many of whom were previously very senior in their own state energy offices during 2008—and a select group of nine states to compliment the states represented by the NASEO Regional Coordinators themselves. KEMA also verified the status of all PAs with the States' Energy Offices, such as whether funding amounts changed, the PA was dropped, and in some cases whether the BPAC assignment had changed.

2.2.2 Decision Rules for Classifying Programmatic Activities

KEMA used the same principles learned from the ARRA data frame development process that is articulated above for assigning BPACs by programmatic activity:

- Assign the BPAC that most fits the programmatic activity.
- Assign the highest level rigor possible that reasonably fits the programmatic activity.
- Assign a secondary or tertiary BPAC if a programmatic activity exhibits such strong supporting elements for further review and disaggregation if additional data were available through the State's Energy Office.
- Assign "Administration" as a BPAC for funded activities that are primarily administrative in nature and have no programmatic feature that would deliver energy savings.
- Assign "Energy Emergency Planning" as a BPAC since the 2008 SEP funding included such a requirement that the ARRA funding did not.

The PY2008 data did not result in any further redefinition or refinement of the BPACs distinguishing attributes.

2.3 Sub-categorization of BPACs

As required in the RFP, the contractor team subcategorized the BPACs as well. As stated in the Final SEP Evaluation White Paper:

“The results from this effort will be the development of a set of program evaluation groups with a description of the characteristics that make the group suitable for grouping and descriptive information about the characteristics of each program that need to feed the efforts for prioritizing programs within an evaluation group.”¹²

Upon review of the PA data, the contractor team determined that not only do the PAs within BPACs disaggregate into subcategories, but also the subcategories may overlap across BPACs as well. For example, the Loans, Grants and Incentives BPAC is at times hard to distinguish from a building retrofit or renewable energy rebate program. Workshops can be conducted across many BPACs, and building retrofit programs can be delivered through technical assistance or audits. As a result, the contractor team found that further specifying the BPACs to a finer level—such as the delivery mechanism or the targeted sector—became a useful basis for subcategorization. This is consistent with the SEP Evaluation White Paper which states that subcategorization efforts should: “...make sure these efforts reflect the way the programs are operated and to accurately capture the services provided.”¹³

The subcategories, largely grounded in the BPACs, developed for this evaluation effort are largely derived from the BPACs, but can be assigned independently of the parent BPAC depending on the delivery mechanism or program target. These are presented in Figure 4.

¹² Hall, Nick, Paul DeCotis, Marty Kushler, Lori Megdal, and Ed Vine. *An Evaluation Approach for Assessing Program Performance from the State Energy Program*. October 2007, Page 21. [Hereafter: “SEP Evaluation White Paper.”]

¹³ Evaluation White Paper, p. 22.

Figure 4: BPACs and Subcategories

BPAC	Subcategory Derivation
Building Codes and Standards	Building Code Development and Support
	End Use Standards Development and Support
Building Retrofits	Building Retrofits: Nonresidential
	Building Retrofits: Residential
Clean Energy Policy Support	Policy and Market Studies; Legislative Support
Energy Audits: Commercial, Industrial and Agricultural	Energy Audits: Commercial, Industrial and Agricultural ¹⁴
Energy Audits: Residential	Energy Audits: Residential
Energy Efficiency Rating and Labeling	Energy Efficiency Rating and Labeling Development and Support
Government, School and Institutional Procurement	Government, School and Institutional Procurement Support
Industrial Process Efficiency	Industrial Retrofit Support
New Construction and Design	New Construction and Design Assistance
Renewable Energy Market Development	Renewable Energy Market Development: Manufacturing
	Renewable Energy Market Development: Projects
Technical Assistance	Technical Assistance to Building Owners
Traffic Signals	Traffic Signal Retrofits
Transportation	Alternative Fuels, Ride Share and Traffic Optimization
Loans, Grants and Incentives	<i>[Never a subcategory]</i> ¹⁵
Tax Incentives and Credits	<i>[Never a subcategory]</i>
<i>[Never a BPAC]</i>	Generalized Marketing and Outreach (Participants not traceable)
<i>[Never a BPAC]</i>	Generalized Workshops and Demonstrations

¹⁴ Details of the sampling strategy are provided in Section 3. PAs with energy audits to the nonresidential sector did not exceed the 3% threshold and were not included for any policy reason, therefore further sub-stratification was unnecessary.

¹⁵ The following BPACs, Loans, Grants and Incentives and Tax Incentives and Credits, are never subcategories, following the White Paper suggestion to ensure that, for the purposes of gross savings estimation, classification efforts, “reflect the way the programs are operated and to accurately capture the services provided.”

BPAC	Subcategory Derivation
	(Participants may be traceable)
<i>[Never a BPAC]</i>	Targeted Training and/or Certification (Participants are traceable)

Additionally, the subcategories were also specified to be consistent with known gross savings estimation methods, such that estimated energy savings by BPAC can be reasonably reflected as the sum of all subcategories.

For each PA, the contractor team assigned a unique BPAC/Subcategory combination to effectively define the sample frame and prioritize. In some cases, this required splitting the record in the PAGE or WinSaga database after verifying the PA's funding level and intent to address the guidance provided in the SEP Evaluation White Paper for prioritization and documentation of design detail.

3. Sampling of Programmatic Activities and Expansion of Sample Results

3.1 Overview of Sampling Approach

The selection of PAs for sampling requires:

1. Establishment of the total sample size,
2. Establishment of the sampling frame, including classification of PAs into groups,
3. A rule or process for assigning the evaluation rigor level to sampled PAs, and
4. A process for allocating sampling points to the groups.

The approach can be summarized as follows.

3.1.1 Total Sample Size

The total number of PAs to be evaluated was set at 82, including 24 High-rigor and 58 Medium-High-rigor PAs, and a total sample size of 53 for PY2008 and 29 for ARRA. These numbers were determined based on an initial assessment of the distribution of funding by activity types, and the number of different types of evaluations that could be accommodated by the available budget.

3.1.2 Sampling Frame

The sampling frame for each period started as the largest BPAC-subcategory cells (in terms of program budget), that together account for at least and not a lot more than 80 percent of non-administrative budget. That is, we defined a minimum funding PA size threshold such that the cells total above but close to 80 percent of the total program budget. All these cells are included in the sampling frame. A few additional cells were then included for policy reasons despite being smaller than the size threshold. The included cells define the population that will be represented by the study.

3.1.3 Rigor Level

After reviewing the activities in the course of the classification process, and in light of budget constraints, we determined that High-rigor evaluations would be meaningful only for evaluation of building retrofit activities. These activities fall into two BPACS: (1) Building Retrofit and (2)

Loans, Grants and Incentives. Under each of these BPACS, there are Residential and Nonresidential building retrofit subcategories. These subcategories are assigned to High-rigor evaluation. All other cells are assigned Medium-High rigor.

3.1.4 PA Sample Allocation

Sample allocation to BPAC-subcategory cells occurs in a few steps.

1. **Preliminary allocation.** Initially PAs are allocated to cells proportional to budget only. This process tends to leave smaller cells, especially those included despite being below the minimum size threshold, with zero allocation.
2. **Forced allocations.** After reviewing the initial allocation strictly proportional to budget, some forced allocations are specified, to ensure the small cells that need to be covered have some sample.
3. **Proportional allocation.** The cells that received forced allocations are set aside. The remainder of the total sample points for each period are allocated to the remaining cells proportional to size (program budget).
4. **Identification of certainty and non-certainty PAs.** Allocation proportional to size means that one sample PA is allocated for about every \$850,000 of budget for PY2008, and for every \$77 million of budget for ARRA.
 - a. Any individual PA with budget above this amount is included with certainty in the sample. The PAs so selected are called “first-pass certainty” PAs. In some cases, the budget for an individual PA would mean an allocation of two or more PAs. However, we only select a given PA once.
 - b. Once the large, first-pass certainty PAs have been identified, the remaining sample points are allocated to the remaining cells, proportional to the remaining size.
 - c. We identify a second set of certainty selections within this remainder sample, using the same approach as for the first pass. That is, all PAs with budget greater than the ratio of total remaining budget to remaining sample size are included with certainty. The PAs so selected are called “second-pass certainty” PAs.
 - d. Once the first- and second-pass certainty PAs have been identified, the remaining sample points are allocated to the remaining cells, proportional to the remaining size. These allocations are referred to as the “non-certainty” or “remainder” sample.

- e. In principle, proportional to size allocation could result in a target sample size greater than the number of PAs in a cell. In such cases it would be necessary to cap the allocations at the number that exist in the cell, and re-allocate the excess sample. This problem of over-allocation did not arise for this sample, so this step was not necessary. Pulling out the certainty cells in two passes helps to reduce the potential for the problem.
5. **Assessment of achievability.** Once we identified the target numbers of certainty and non-certainty selections for each cell, we assessed whether there are cells whose targets are unlikely to be met based on evaluability.
- a. Each PA has an evaluability score indicating either a high or moderate chance of successfully completing an evaluation at the targeted rigor if we select that PA. (PAs with zero or low chance of successful Medium-High- or High-rigor evaluation account for a small fraction of total activity, and are excluded from the frame.) Specifically, we assume that a “likely” evaluable PA has an 80 percent chance of being evaluated at the targeted High or Medium-High rigor level, while a “possibly” evaluable PA has a 50 percent chance. Based on discussions with representatives from DOE, ORNL and the states who participated in the May 25th Network Committee Meeting, we feel that these are conservative estimates.
 - b. The assumed success rates should be very conservative for certainty PAs. Certainty PAs are high priority for successful completion because of their size. If after confirming with ORNL that we are unable to complete evaluation of one of these PAs, we will substitute a smaller PA. However, this substitution will be a last resort.
 - c. The remainder sample will be allocated to “likely evaluable” PAs at a higher rate than “possibly evaluable” PAs. This procedure ensures that both levels of evaluability are covered by the sample, but that evaluation resources are devoted more heavily to the PAs that have better chance of being evaluable.
 - d. Based on the assumed probabilities of successful evaluation at targeted rigor for likely and possible, we calculate the size of the oversample required to achieve the targeted sample sizes. With the assumed success probabilities, we need a sample of five “likely” PAs to complete four evaluations successfully. We need a sample of two “possible” PAs to complete one evaluation successfully.
 - e. If the total oversample required based on this calculation exceeds the number of PAs in the sample, we flag a potential shortfall. As it turned out, the current sample design does not have an anticipated shortfall in any cell. That is, unless the frequency of inadequate data availability is worse than projected in some cell, we expect to achieve these targeted sample sizes at the targeted rigor levels.

6. **Final targets.** After the iterative reallocation in steps 1-4, we reviewed the sample allocations and made some slight adjustments to be sure:
 - a. Total samples after rounding still matched the targeted number by time period and by rigor level, and
 - b. The iterative re-allocation of the remainder did not result in severe over- or under-allocation to any one cell.

The remainder of this section presents the results of each of these steps.

3.2 Sample Frame

Figure 5 indicates the proportion of SEP spending for each BPAC and Subcategory for PY2008 and the Figure 6 presents similar data for the ARRA period. As noted, our starting point for frame definition is to select the BPAC/Subcategory combinations that sum to at least 80 percent of funding. The minimum funding percentage by BPAC/Subcategory combination is 3 percent for both periods (pink highlighted cells). In addition, we have included select BPAC/Subcategory combinations (yellow highlighted cells) that may be outside the sampling criteria for policy reasons to ensure adequate inclusion of important BPACs. The additional included cells are the following:

- Building Codes and Standards—this BPAC is anticipated to produce savings disproportionate to spending.
- Subcategories of Workshops/Demonstrations and Training/Certification that are likely to be evaluable, if the other subcategories of the associated BPAC are included.
- Building Retrofit subcategories if not already included based on size.

As shown in Figure 5, the sampling approach represents 80.3 percent of SEP funding for PY2008, and as shown in Figure 6 the sampling approach represents 86.4 percent for the ARRA period. Figure 7 and Figure 8 display the number of available PAs within each of the selected BPAC/subcategory combinations for PY2008 and the ARRA period, respectively. Pink cells represent PA BPAC/subcategory combinations which exceed the 3 percent minimum threshold; yellow cells are those BPAC/subcategory combinations which are included for policy reasons. As shown, 173 PAs are included in the sampling frame for PY2008 and 355 are included in the sampling frame for the ARRA period.

Figure 5: Percent of PY2008 SEP Budget by BPAC and Subcategory

BPACs	Subcategory																All Subcategories	Selected BPACs/ Subcategories
	Alternative Fuels, Ride Share and Traffic Optimization	Building Code Development and Support	Building Retrofits: Nonresidential	Building Retrofits: Residential	Energy Audits: Commercial, Industrial and Agricultural	Energy Audits: Residential	Generalized Marketing and Outreach	Generalized Workshops and Demonstrations	Government, School and Institutional Procurement	Industrial Retrofit Support	New Construction and Design	Policy and Market Studies ; Legislative Support	Renewable Energy Market Development: Manufacturing	Renewable Energy Market Development: Projects	Targeted Training and/or Certification	Technical Assistance to Building Owners		
Building Codes and Standards	0.0%	0.9%	0.0%	0.0%	0.0%	0.0%	0.0%	3.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.6%	5.2%	10.7%	10.7%
Building Retrofits	0.0%	0.0%	1.6%	1.0%	0.5%	0.1%	2.3%	5.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.3%	7.2%	18.1%	12.2%
Clean Energy Policy Support	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	10.1%	0.0%	0.0%	0.0%	0.0%	10.1%	10.1%
Energy Audits: Commercial, Industrial and Agricultural	0.0%	0.0%	0.0%	0.0%	2.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	2.4%	
Energy Audits: Residential	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.6%	
Energy Efficiency Rating and Labeling	0.0%	0.9%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.1%	
Government, School and Institutional Procurement	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.2%	
Industrial Retrofit Support	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.5%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.7%	
Loans, Grants and Incentives (Excl Retro)	5.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.7%	0.2%	0.0%	0.0%	0.0%	0.0%	0.5%	0.7%	0.0%	8.9%	16.2%	14.2%

Figure 5: Percent of PY2008 SEP Budget by BPAC and Subcategory

BPACs	Subcategory																All Subcategories	Selected BPACs/ Subcategories
	Alternative Fuels, Ride Share and Traffic Optimization	Building Code Development and Support	Building Retrofits: Nonresidential	Building Retrofits: Residential	Energy Audits: Commercial, Industrial and Ag	Energy Audits: Residential	Generalized Marketing and Outreach	Generalized Workshops and Demonstrations	Government, School and Institutional Procurement	Industrial Retrofit Support	New Construction and Design	Policy and Market Studies ; Legislative Support	Renewable Energy Market Development: Manufacturing	Renewable Energy Market Development: Projects	Targeted Training and/or Certification	Technical Assistance to Building Owners		
Loans, Grants and Incentives (Retro Only)	0.0%	0.0%	16.5%	8.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	24.6%	24.6%
New Construction and Design	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.8%	
Renewable Energy Market Development	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	5.1%	0.0%	0.0%	0.0%	0.0%	0.0%	1.6%	0.0%	0.0%	7.2%	5.1%
Tax Incentives and Credits	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.6%	0.0%	0.6%	
Technical Assistance	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	3.2%	3.5%	3.3%
Transportation	0.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.8%	1.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	3.2%	
All BPACs	5.6%	1.8%	18.1%	9.1%	2.8%	0.3%	4.9%	16.0%	0.0%	0.1%	0.0%	10.1%	0.5%	2.4%	3.6%	24.5%	100.0%	
Selected BPACs/ Subcategories	5.2%	0.9%	16.5%	8.1%	0.0%	0.0%	0.0%	13.5%	0.0%	0.0%	0.0%	10.1%	0.0%	0.0%	1.6%	24.5%		80.3%

Figure 6: Percent of ARRA SEP Budget by BPAC and Subcategory

BPACs	Subcategory																		All Subcategories	Selected BPACs/Subcategories
	Alternative Fuels, Ride Share and Traffic Optimization	Building Codes and Standards: Codes	Building Retrofits: Nonresidential	Building Retrofits: Residential	Energy Audits: Commercial, Industrial and Ag	Energy Audits: Residential	Energy Efficiency Rating and Labeling	Generalized Marketing and Outreach	Generalized Workshops and Demonstrations	Government, School and Institutional Procurement	Industrial Retrofit Support	New Construction and Design	Policy and Market Studies; Legislative Support	Renewable Energy Market Development: Manufacturing	Renewable Energy Market Development: Projects	Targeted Training and/or Certification	Technical Assistance to Building Owners	Traffic Signals		
Building Codes and Standards	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	1.3%	1.3%
Building Retrofits	0.0%	0.0%	22.9%	2.7%	0.8%	0.2%	0.0%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.0%	0.5%	0.0%	28.5%	24.0%
Clean Energy Policy Support	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.0%	
Energy Audits: Commercial, Industrial and Agricultural	0.0%	0.0%	0.0%	0.0%	0.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.7%	
Energy Audits: Residential	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	
Energy Efficiency Rating and Labeling	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	
Industrial Retrofit Support	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.9%	
Loans, Grants and Incentives (Excl Retrofits and Projects)	0.8%	0.0%	0.0%	0.0%	0.3%	0.0%	0.1%	0.0%	0.2%	0.0%	1.6%	0.0%	0.0%	8.4%	0.0%	0.4%	0.2%	0.0%	11.9%	8.9%

Figure 6: Percent of ARRA SEP Budget by BPAC and Subcategory

BPACs	Subcategory																		All Subcategories	Selected BPACs/Subcategories
	Alternative Fuels, Ride Share and Traffic Optimization	Building Codes and Standards: Codes	Building Retrofits: Nonresidential	Building Retrofits: Residential	Energy Audits: Commercial, Industrial and Ag	Energy Audits: Residential	Energy Efficiency Rating and Labeling	Generalized Marketing and Outreach	Generalized Workshops and Demonstrations	Government, School and Institutional Procurement	Industrial Retrofit Support	New Construction and Design	Policy and Market Studies; Legislative Support	Renewable Energy Market Development: Manufacturing	Renewable Energy Market Development: Projects	Targeted Training and/or Certification	Technical Assistance to Building Owners	Traffic Signals		
Loans, Grants and Incentives (Retrofits and Projects)	0.0%	0.0%	18.6%	5.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	11.6%	0.0%	0.0%	0.0%	35.3%	35.3%
New Construction and Design	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.4%	
Renewable Energy Market Development	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	4.7%	11.7%	0.6%	0.2%	0.0%	17.2%	16.9%	
Technical Assistance	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.3%	0.0%	0.3%	
Traffic Signals	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.2%	
Transportation	2.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.1%	
All BPACs	2.8%	0.5%	41.5%	7.8%	1.8%	0.4%	0.1%	0.6%	1.0%	0.0%	2.4%	0.2%	1.0%	13.1%	23.2%	2.2%	1.3%	0.2%	100.0%	
Selected BPAC/Subcategories	0.0%	0.4%	41.5%	5.1%	0.0%	0.0%	0.0%	0.0%	1.0%	0.0%	0.0%	0.0%	0.0%	13.1%	23.2%	2.1%	0.0%	0.0%		86.4%

Figure 7: Number of Available PAs for Selected BPAC/Subcategory (PY2008)

BPAC	Subcategory	Target Rigor	SEP Budget	Percent of SEP Budget	Number of PA's
Building Codes and Standards	Building Code Development and Support	MH	\$507,271	1%	5
Building Codes and Standards	Generalized Workshops and Demonstrations	MH	\$1,735,812	4%	7
Building Codes and Standards	Targeted Training and/or Certification	MH	\$882,531	2%	6
Building Codes and Standards	Technical Assistance to Building Owners	MH	\$2,972,522	6%	3
Building Retrofits	Building Retrofits: Nonresidential	H	\$913,228	2%	7
Building Retrofits	Building Retrofits: Residential	H	\$576,183	1%	5
Building Retrofits	Generalized Workshops and Demonstrations	MH	\$2,862,000	6%	24
Building Retrofits	Targeted Training and/or Certification	MH	\$191,447	0%	6
Building Retrofits	Technical Assistance to Building Owners	MH	\$4,074,048	9%	13
Clean Energy Policy Support	Policy and Market Studies; Legislative Support	MH	\$5,714,771	12%	39
Loans, Grants and Incentives (Excl Retro)	Alternative Fuels, Ride Share and Traffic Optimization	MH	\$2,932,203	6%	10
Loans, Grants and Incentives (Excl Retro)	Generalized Workshops and Demonstrations	MH	\$97,222	0%	3
Loans, Grants and Incentives (Excl Retro)	Technical Assistance to Building Owners	MH	\$5,062,979	11%	4
Loans, Grants and Incentives (Retro Only)	Building Retrofits: Nonresidential	H	\$9,392,550	20%	8
Loans, Grants and Incentives (Retro Only)	Building Retrofits: Residential	H	\$4,614,510	10%	3

Figure 7: Number of Available PAs for Selected BPAC/Subcategory (PY2008)					
BPAC	Subcategory	Target Rigor	SEP Budget	Percent of SEP Budget	Number of PA's
Renewable Energy Market Development	Generalized Workshops and Demonstrations	MH	\$2,926,128	6%	19
Technical Assistance	Generalized Workshops and Demonstrations	MH	\$75,402	0%	3
Technical Assistance	Targeted Training and/or Certification	MH	\$95,880	0%	1
Technical Assistance	Technical Assistance to Building Owners	MH	\$1,801,193	4%	7
			\$47,427,881	100%	173

Figure 8: Number of Available PAs for Selected BPAC/Subcategory (ARRA)

BPAC	Subcategory	Target Rigor	SEP Budget	Percent of SEP Budget	Number of PA's
Building Codes and Standards	Building Codes and Standards: Codes	MH	\$11,356,748	0%	15
Building Codes and Standards	Generalized Workshops and Demonstrations	MH	\$19,223,610	1%	2
Building Codes and Standards	Targeted Training and/or Certification	MH	\$2,489,921	0%	10
Building Retrofits	Building Retrofits: Nonresidential	H	\$585,731,006	26%	86
Building Retrofits	Building Retrofits: Residential	H	\$69,377,772	3%	16
Building Retrofits	Generalized Workshops and Demonstrations	MH	\$667,990	0%	1
Building Retrofits	Targeted Training and/or Certification	MH	\$26,537,692	1%	11
Loans, Grants and Incentives (Excl Retrofits and Projects)	Generalized Workshops and Demonstrations	MH	\$4,047,962	0%	3
Loans, Grants and Incentives (Excl Retrofits and Projects)	Renewable Energy Market Development: Manufacturing	MH	\$216,947,443	9%	9
Loans, Grants and Incentives (Excl Retrofits and Projects)	Targeted Training and/or Certification	MH	\$9,558,163	0%	3
Loans, Grants and Incentives (Retrofits and Projects)	Building Retrofits: Nonresidential	H	\$479,418,126	21%	45
Loans, Grants and Incentives (Retrofits and Projects)	Building Retrofits: Residential	H	\$135,981,963	6%	16
Loans, Grants and Incentives (Retrofits and Projects)	Renewable Energy Market Development: Projects	MH	\$295,725,557	13%	57

Figure 8: Number of Available PAs for Selected BPAC/Subcategory (ARRA)

BPAC	Subcategory	Target Rigor	SEP Budget	Percent of SEP Budget	Number of PA's
Renewable Energy Market Development	Generalized Workshops and Demonstrations	MH	\$1,108,465	0%	5
Renewable Energy Market Development	Renewable Energy Market Development: Manufacturing	MH	\$120,323,694	5%	10
Renewable Energy Market Development	Renewable Energy Market Development: Projects	MH	\$299,531,840	13%	58
Renewable Energy Market Development	Targeted Training and/or Certification	MH	\$14,852,017	1%	8
			\$2,292,879,968	100%	355

3.3 Sampling Targets

Our overall sampling targets were established based on the desired level of effort and available resources. The total number of PAs to be sampled was set at 53 for PY2008 and 29 for the ARRA period. Preliminary rigor level assignments were specified as follows:

- PY2008: 24 High-rigor evaluations, 29 Medium-high-rigor evaluations
- ARRA: 29 Medium-high-rigor evaluations

However, as noted, we determined that given the results of the activity classifications and in light of budget constraints, only a limited number of cells were amenable to High-rigor evaluation. Limiting High-rigor evaluation to PY2008, while retaining the target of 24, would heavily direct the PY2008 sample to only a few types of activities. Instead, we plan to distribute the High-rigor evaluations between PY2008 and ARRA.

Within these overall guidelines, we followed the steps outlined in Section 3.1.4 above. This allocation resulted in the sampling targets shown in figure 9 and Figure 10 for the PY2008 and ARRA periods, respectively.

The figures show the allocation that would be assigned based strictly by allocating proportional to total budget (green highlighted columns), and also the allocations that would result from allocating strictly proportional to the number of PAs in the cell (red highlighted cells). Also shown is the total number allocated through the iterative process in the blue highlighted cells, combining the certainty and non-certainty PAs.

The figures show a few cells with allocations of zero. These are cells initially included in the frame, but that were too small to receive an allocation with proportional allocation. These were all cells that were included in the frame to ensure some coverage of evaluable Workshops/Demonstrations and Training/Certification (subcategory) activities. We did not force allocations to these cells, because enough other activities in these subcategories were included.

There are a few cells (highlighted in yellow) where the final proposed allocation differs from the iteratively allocated targets (in blue).

- For PY2008, the iterative allocation results in a target of 10 for Clean Energy Policy Support. This allocation would be 19 percent of the sample, for 12 percent of the budget and 23 percent of the number of PAs. We reduced this allocation to eight, and added one each to Building Retrofit/Technical Assistance to Building Owners and to Renewable

Energy Market Development/Generalized Workshops and Demonstrations (yellow highlighted cells).

- For ARRA, the rounding of cell targets resulted in a total of 27 selections instead of the targeted 29. We added one each to Loans, Grants and Incentives/Renewable Energy Market Development: Projects and Renewable Energy Market Development/Renewable Energy Market Development: Projects (yellow highlighted cells).

The figures also show that in most cases the proposed targets are within the range bracketed by allocation proportional to size and allocation proportional to number of PAs. Allocations less than proportional to size are mostly associated with large numbers of certainty selections.

Finally, the figures indicate that the targets are expected to be achievable based on the numbers available in each cell and the assumed success rates. That is, the likely shortfall is zero.



Figure 9: Allocated Sampling Targets by BPAC/Subcategory and Rigor Level (PY2008)

BPAC	Sub-category	Target Rigor	Budget			Population # PAs			Iteratively Allocated Sample Size				Likely Shortfall	Final Proposed Target
			Budget	% Budget	Sample Proportional to Budget	Population # PAs	% Population # PAs	Sample Proportional to # PAs	Certainty	Non-Certainty	Total	% Sample Total		
Building Codes and Standards	Building Code Development and Support	MH	\$507,271	1%	1	5	3%	2	0	1	1	2%	0	1
Building Codes and Standards	Generalized Workshops and Demonstrations	MH	\$1,735,812	4%	2	7	4%	2	1	2	3	6%	0	3
Building Codes and Standards	Targeted Training and/or Certification	MH	\$882,531	2%	1	6	3%	2	0	2	2	3%	0	2
Building Codes and Standards	Technical Assistance to Building Owners	MH	\$2,972,522	6%	3	3	2%	1	1	0	1	2%	0	1
Building Retrofits	Building Retrofits: Nonresidential	H	\$913,228	2%	1	7	4%	2	0	2	2	4%	0	2
Building Retrofits	Building Retrofits: Residential	H	\$576,183	1%	1	5	3%	2	0	2	2	4%	0	2
Building Retrofits	Generalized Workshops and Demonstrations	MH	\$2,862,000	6%	3	24	14%	7	1	4	5	9%	0	5
Building Retrofits	Targeted Training and/or Certification	MH	\$191,447	0%	0	6	3%	2	0	0	0	1%	0	0
Building Retrofits	Technical Assistance to Building Owners	MH	\$4,074,048	9%	5	13	8%	4	1	4	5	10%	0	6
Clean Energy Policy Support	Policy and Market Studies; Legislative Support	MH	\$5,714,771	12%	6	39	23%	12	2	8	10	19%	0	8
Loans, Grants and Incentives (Excl Retro)	Alternative Fuels, Ride Share and Traffic Optimization	MH	\$2,932,203	6%	3	10	6%	3	1	4	5	9%	0	5
Loans, Grants and Incentives (Excl Retro)	Generalized Workshops and Demonstrations	MH	\$97,222	0%	0	3	2%	1	0	0	0	0%	0	0
Loans, Grants and Incentives (Excl Retro)	Technical Assistance to Building Owners	MH	\$5,062,979	11%	6	4	2%	1	2	1	3	5%	0	3
Loans, Grants and Incentives (Retro Only)	Building Retrofits: Nonresidential	H	\$9,392,550	20%	10	8	5%	2	3	1	4	8%	0	4
Loans, Grants and Incentives (Retro Only)	Building Retrofits: Residential	H	\$4,614,510	10%	5	3	2%	1	2	0	2	4%	0	2
Renewable Energy Market Development	Generalized Workshops and Demonstrations	MH	\$2,926,128	6%	3	19	11%	6	2	3	5	9%	0	6
Technical Assistance	Generalized Workshops and Demonstrations	MH	\$75,402	0%	0	3	2%	1	0	0	0	0%	0	0
Technical Assistance	Targeted Training and/or Certification	MH	\$95,880	0%	0	1	1%	0	0	0	0	0%	0	0
Technical Assistance	Technical Assistance to Building Owners	MH	\$1,801,193	4%	2	7	4%	2	2	1	3	5%	0	3
		Total	\$47,427,881	100%	53	173	100%	53	18	35	53	100%	0	53
		MH	\$1,931,410	67%	36	150	87%	46	13	30	43	81%	0	43

Figure 9: Allocated Sampling Targets by BPAC/Subcategory and Rigor Level (PY2008)

BPAC	Sub-category	Target Rigor	Budget			Population # PAs			Iteratively Allocated Sample Size				Likely Shortfall	Final Proposed Target
			Budget	% Budget	Sample Proportional to Budget	Population # PAs	% Population # PAs	Sample Proportional to # PAs	Certainty	Non-Certainty	Total	% Sample Total		
		H	15,496,471	33%	17	23	13%	7	5	5	10	19%	0	10



Figure 10: Allocated Sampling Targets by BPAC/Subcategory and Rigor Level (ARRA)

BPAC	Sub-category	Target Rigor	Budget			Population # PAs			Iteratively Allocated Sample Size				Likely Shortfall	Final Proposed Target
			Budget	% Budget	Sample Proportional to Budget	Population # PAs	% Population # PAs	Sample Proportional to # PAs	Certainty	Non-Certainty	Total	% Sample Total		
Building Codes and Standards	Building Codes and Standards: Codes	MH	\$11,356,748	0%	0	15	4%	1	0	2	2.0	7%	0	2
Building Codes and Standards	Generalized Workshops and Demonstrations	MH	\$19,223,610	1%	0	2	1%	0	0	1	1.0	3%	0	1
Building Codes and Standards	Targeted Training and/or Certification	MH	\$2,489,921	0%	0	10	3%	1	0	1	1.0	3%	0	1
Building Retrofits	Building Retrofits: Nonresidential	H	\$585,731,006	26%	7	86	24%	7	0	6	6.3	22%	0	6
Building Retrofits	Building Retrofits: Residential	H	\$69,377,772	3%	1	16	5%	1	0	2	2.0	7%	0	2
Building Retrofits	Generalized Workshops and Demonstrations	MH	\$667,990	0%	0	1	0%	0	0	0	0.0	0%	0	0
Building Retrofits	Targeted Training and/or Certification	MH	\$26,537,692	1%	0	11	3%	1	0	0	0.3	1%	0	0
Loans, Grants and Incentives (Excl Retrofits and Projects)	Generalized Workshops and Demonstrations	MH	\$4,047,962	0%	0	3	1%	0	0	0	0.0	0%	0	0
Loans, Grants and Incentives (Excl Retrofits and Projects)	Renewable Energy Market Development: Manufacturing	MH	\$216,947,443	9%	3	9	3%	1	0	2	2.3	8%	0	2
Loans, Grants and Incentives (Excl Retrofits and Projects)	Targeted Training and/or Certification	MH	\$9,558,163	0%	0	3	1%	0	0	0	0.1	0%	0	0
Loans, Grants and Incentives (Retrofits and Projects)	Building Retrofits: Nonresidential	H	\$479,418,126	21%	6	45	13%	4	1	4	4.7	16%	0	5
Loans, Grants and Incentives (Retrofits and Projects)	Building Retrofits: Residential	H	\$135,981,963	6%	2	16	5%	1	0	1	1.5	5%	0	1
Loans, Grants and Incentives (Retrofits and Projects)	Renewable Energy Market Development: Projects	MH	\$295,725,557	13%	4	57	16%	5	0	3	3.2	11%	0	4
Renewable Energy Market Development	Generalized Workshops and Demonstrations	MH	\$1,108,465	0%	0	5	1%	0	0	0	0.0	0%	0	0
Renewable Energy Market Development	Renewable Energy Market Development: Manufacturing	MH	\$120,323,694	5%	2	10	3%	1	0	1	1.3	4%	0	1
Renewable Energy Market Development	Renewable Energy Market Development: Projects	MH	\$299,531,840	13%	4	58	16%	5	0	3	3.2	11%	0	4

Figure 10: Allocated Sampling Targets by BPAC/Subcategory and Rigor Level (ARRA)

BPAC	Sub-category	Target Rigor	Budget			Population # PAs			Iteratively Allocated Sample Size				Likely Shortfall	Final Proposed Target
			Budget	% Budget	Sample Proportional to Budget	Population # PAs	% Population # PAs	Sample Proportional to # PAs	Certainty	Non-Certainty	Total	% Sample Total		
Renewable Energy Market Development	Targeted Training and/or Certification	MH	\$14,852,017	1%	0	8	2%	1	0	0	0.2	1%	0	0
		Total	\$2,292,879,968	100%	29	355	100%	29	1	28	29	100%	0	29
		MH	\$1,022,371,101	45%	13	192	54%	16	0	15	15	50%	0	15
		H	\$1,270,508,867	55%	16	163	46%	13	1	13	14	50%	0	14

3.4 Implementing the PA Sample Design

Drawing the PA sample design means selecting a simple random sample from each sampling cell, with the cell sample size specified by the design. For each selected PA, the plan is to conduct an evaluation at the Rigor level specified by the design. As discussed above, there may be some obstacles to conducting the evaluations at the desired rigor levels.

3.4.1 Misclassification and Multiple Classifications

Detailed investigation of a selected PA may reveal that it has been incorrectly classified at the frame development stage. In addition, many PAs are known even from the currently available data to include multiple categories of activity.

To deal with both misclassification and multiple categories, we distinguish between the *sampling* category and the *analytic* category or reporting domain. PAs are assigned to BPACs at the sample design and frame development stage based on the information available from the data bases. This assignment and the sample allocation determine each PA's probability of being included in the sample. That probability determines its sample expansion weight, and its stratum assignment for the calculation of ratio estimates and standard errors.

For purposes of *analysis*, activities may be classified by information available at the design stage, or by information available only after collecting more information from the selected PAs. Information can be reported for all components of all PAs that include a certain type of activity, not just for the PAs assigned to a particular category for sampling. Thus, for example, to determine the total savings from all residential retrofits, as identified post-sampling, we would sum up the residential retrofit components in all sampling strata, each weighted by that stratum's expansion weight.

This situation is analogous to stratifying buildings based on imperfect building type information. Each building may have multiple types of activities. A sample is stratified based on the best information available at the sampling stage to classify buildings by predominant activity type. During data collection, information may be obtained about the portions of the building corresponding to each activity type. Information can then be reported by domains corresponding to observed activity types. The weighting and stratification are based on the sampling information.

3.4.2 Independent State-Specific Evaluations

Some states are currently conducting their own evaluations of ARRA SEP. The national evaluation needs to apply consistent methods across all states, and to ensure the quality of the work. At the same time, to the extent that the state-level evaluations do provide reliable results, incorporating these results into the national estimates can improve these estimates. This is particularly true since the states with individual evaluations tend to be among the larger recipients of ARRA SEP funding, and the state-level evaluations in total are of comparable magnitude to this national ARRA SEP evaluation. Thus, it is important to consider whether and how the national evaluation can incorporate information from the state evaluations, as well as how the two efforts can be coordinated so that both are successful.

If the individual state evaluations were all evaluated using the same methods as this national evaluation, state evaluation results could be incorporated directly into the national evaluation, with adjustments to the sample design. To illustrate, suppose that a state evaluation provides an estimate consistent with our methodology for a group of PAs within that state's SEP portfolio. In that case, the PAs evaluated by their own state would be in the sample representing themselves, and the random sample would represent the PAs from other states. The state-evaluated PAs would have expansion weights of one (they represent only themselves). The other PAs in the sample would have expansion weights larger than one (they represent themselves and others), but smaller than if the state sample were not incorporated.

With this weighting procedure, combining the state and national samples does not bias the national results. The approach produces results with higher accuracy than if all states were represented by the national random sample. This method of combining two samples is fully consistent with statistical sampling principles and is used regularly by federal agencies including DOE's Energy Information Administration.

In practice, we do not expect the state evaluations to follow identical procedures to those for this national evaluation. Nevertheless there are ways this evaluation can use the state-level efforts.

For sampling, we will first understand as well as possible what the state evaluations are doing. If the state evaluation is using methods that are fully consistent with this evaluation's rigor levels, methods, and assumptions, we can use any results they develop, as indicated above.

The situation is more complex if the rigor and methods of the state evaluation are consistent with what we would consider acceptable, but their assumption are different (in particular,

baseline definitions and measure life). In such cases, we will request access to the state evaluation's analysis data sets and adjust these key assumptions to develop adjusted savings consistent with ours.

If the state evaluation methods and rigor are acceptable, adjusting their results for consistency of assumptions should be a lower cost effort compared to doing primary data collection and analysis. This adjustment process does still have costs. Thus, we would not automatically incorporate all state results even if there are no methodological issues. We will prioritize those that provide good information for large portions of the BPACs targeted by this national evaluation.

If we determine that a state's methods are too inconsistent with the national evaluation methods to be incorporated even with adjustments, the PAs from that state will be included or not in the national evaluation sample based on the random selection, just as if there were no state evaluation. If we select a PA where we know the state has already done work, we will attempt to take advantage of the data sets compiled for the state evaluation. We will also try to avoid re-sampling the same individuals if possible. To the extent that it is necessary for our evaluation to interview some of the same people as the state evaluation, we will try to coordinate with the state evaluation on this data gathering. Depending on the timing, we may make use of information already collected by the state evaluation, and do callbacks to get additional information as needed.

This coordination issue arises only for ARRA, as the state evaluations are not addressing the pre-ARRA period. Whatever we do will require coordination with the state evaluators, which in turn will require support from the sponsoring state's SEP office.

3.5 BPAC-Specific Impact Calculations

For each selected PA, our evaluation will produce calculated impacts and error bounds. We will also have one or more measure of size (MOS) for each PA. At a minimum we will have the spending amount. We also may have more informative correlates of savings such as program-estimated impacts, or other activity measure such as number of units or square feet affected.

From these results we will calculate a statistical ratio estimate for the BPAC for each of the key metrics estimated from the PA sample. We will use the Combined rather than Stratified form of the ratio estimator, because the latter form has more bias when stratum sample sizes are small

as in this study. Specifically, we calculate the ratio R^{\wedge} of the stratified estimate of population savings (or other metric) to the corresponding estimate of population measure of size from the same sample:

$$R^{\wedge} = \sum_k N_k y_{\cdot k} / \sum_k N_k x_{\cdot k}$$

Where

N_k = population number of individual PAs in stratum k

$y_{\cdot k}$ = sample mean savings for stratum k

$x_{\cdot k}$ = sample mean MOS for stratum k.

This ratio is a form of unit savings estimate. For example, if the measure of size x is the number of square feet audited, the ratio R^{\wedge} is savings per square foot audited. We then calculate the population total savings Y_{TOT} by multiplying the total measure of size X_{TOT} known from the data base by this ratio (e.g., multiply savings per square foot by total square feet audited):

$$Y_{TOT} = R^{\wedge} X_{TOT}.$$

We will calculate the standard error of the ratio and the corresponding total savings estimate via statistical formulas for stratified ratio estimation, e.g. from Cochran (1971).

Since there will be few observations in each ratio, we will need to consider if all of these are equally reliable, or if some should be down-weighted in developing the combined ratio. We may also use post-stratification if a selection appears to be substantially atypical.

3.5.1 Portfolio-Level Impact Calculations

The procedures described above will provide estimates of savings and other impacts for each BPAC. Total impacts for PAs represented by 80.3 percent of funding in 2008 PAs and 86.4 percent of funding for ARRA will be calculated as the sum of the impacts by BPAC for each program year. However, as noted, some parameters determined from one period at a higher rigor may be used in calculations for the other period.

3.5.2 Error Bounds

The BPAC-level evaluations are designed to be conducted at different levels of rigor. For the High- and Medium-High-Rigor evaluations, formal accuracy measures will be available from the statistical procedures. For the Medium-Low-rigor evaluations, if any are required, there will be substantial non-statistical uncertainties.

We will provide a discussion of both statistical and non-statistical uncertainties in the BPAC estimates. To the extent possible, we will provide bounds on the non-statistical uncertainties, and combine the statistical and non-statistical uncertainties via error propagation formulas. These formulas will essentially treat the statistical and non-statistical uncertainties as independent sources of variance, and sum independent variance components. The formulas will take into account components of the impact calculation that are appropriately treated as independent and those that rely on common assumptions or statistical samples. Monte Carlo tools may be used to combine the uncertainties and account for the dependencies.

4. Estimation of Energy Impacts

4.1 Introduction

This section provides guidelines for estimating the energy impacts for all sets of programmatic activities to be evaluated, grouped by BPAC and rigor level as described in Section 3. The energy impacts referred to in this section correspond in concept to “gross savings” as that term is commonly used in evaluation of rate payer-funded energy efficiency programs. Evaluation team members charged with managing each of the sample PA evaluations (the Lead Evaluators) will prepare detailed evaluation plans that take into account each sample PA’s actual operations, scale, organization, roster of services provided, and level of documentation.¹⁶

4.1.1 Framework for Specification of Impact Assessment Methods

The following considerations form the framework for the energy impact assessment methods proposed below.

- **Rigor levels and resource allocations.** The assignment of rigor levels carries with it assumptions about the level of resources that will be needed to carry out the individual PA evaluations. Specifically, it was assumed that high-rigor evaluations will require 800 hours on average to complete. Medium-high-rigor evaluations will require an average of 570 hours to complete. Our current sampling plan does not call for implementation of medium-low-rigor evaluations. However, should we need to conduct such evaluations for a given PA due to lack of information or cooperation, we anticipate that medium-low-rigor evaluations will require roughly 200 hours to complete. Our proposed methods are designed to fit within these resource allocations, given the team’s best judgment based on extensive experience in evaluating energy efficiency and renewable energy programs.
- **Information on levels of PA documentation.** Members of the KEMA team have collected information on the organization of PA-level documentation through our initial round of calls and document review, as well as through work on SEP activity evaluations sponsored by state government agencies and other organizations. Based on this experience, we have concluded that, in most cases, considerable effort will be required to organize, compile, and occasionally reconstruct program records, particularly those for PAs sampled in PY2008. Because these records serve as the basis for impact

¹⁶ The content and format of these plans are discussed in Section 8.

assessment at all rigor levels, we have allocated significant resources to the review and improvement of tracking system data quality.

- **Minimization of data collection processes requiring OMB review.** See discussion of this point in Section 1.

4.1.2 Groupings of Programs for Energy Impact Assessment Planning

To expedite the presentation we have grouped the BPACs discussed in Section 3 according to the mechanism by which savings are achieved. These mechanisms strongly shape the kinds of energy impact estimation methods that can be successfully applied at the various rigor levels. Specifically, we have identified the following four basic groups of programmatic activities supported by SEP for development of impact assessment guidelines.

1. **Building Retrofit and Equipment Replacement.** This basic energy savings mechanism involves the implementation of energy-savings capital projects or the installation of energy-efficient equipment in existing residential, commercial, and industrial facilities. Estimation of energy savings generally requires the following steps:
 - a. Review and validation of program records to ensure that they capture and characterize accurately the capital improvements or efficient equipment installations supported by the program.
 - b. Estimation of *ex ante* savings using industry standard engineering methods and input variables drawn from the program records.
 - c. Measurement and verification of the installation and operation of a sample of projects supported by the projects and estimation of energy savings for each project in the sample using the new, verified information.
 - d. Expansion of sample findings to the population of projects, usually through the application of ratio estimation.

In a limited set of cases, other kinds of verification strategies, such as building simulation modeling incorporating various types of data on participating facilities can be used to estimate changes in energy use associated with customer participation in the program. Similarly, the evaluation team may opt to use a billing analysis approach if billing data

can be obtained and other conditions for the application of this family of methods are present. These would include some but not necessarily all of the following: samples of sufficient size, availability of data on non-participants, and ability to obtain releases from individual customers so that billing data can be obtained.

Covered Broad Programmatic Activity Categories and Subcategories. Building Retrofits: Non-Residential and Residential; Loans, grants, and incentives; all subcategories except

- 2. Renewable energy supplied and/or capacity installed.** This energy savings mechanism involves the production and delivery of energy using renewable technologies that would otherwise have been produced by conventional fuels including: petroleum products, natural gas, nuclear power, or coal. Estimation of energy savings for projects that use this type of mechanism typically take an approach similar to measurement and verification of savings building retrofits:
 - a. Review and validation of program records to ensure that they capture and characterize accurately the renewable energy equipment installations made with program support.
 - b. Verification of the installation and operation of a sample of installations supported by the projects including metering of output over a period of time if record reviews are deemed insufficient.
 - c. Re-estimation of energy production for each project in the sample using the new, verified information, as well as routines for annualizing consumption from observed periods.
 - d. Expansion of sample findings to the population of projects, usually through the application of ratio estimation, with appropriate segmentation by renewable energy system type and size.

In some cases, other approaches to estimation of site level savings may be available, including renewable energy system simulation and modeling, and use of longer-term energy production data from meters installed as part of the renewable energy system.

Covered Broad Programmatic Activity Categories. Renewable energy market development – projects; Renewable energy market development – manufacturing; Clean energy policy support.

3. Information and Training Programs. Information and training programs include activities that we have sorted into the following subcategories: Generalized Workshops and Demonstrations, Targeted Technical Training, and Technical Assistance to Building Owners. This group encompasses PAs that include the use of information and, in some cases, custom facility analysis, to motivate and enable customers to undertake energy-efficiency improvements to capital facilities and/or operating practices. Other components of these programs provide direct financial support for the capital and operating improvements. The first step in assessing the energy impacts of such PAs is to determine whether and how participants changed their investment, purchasing, or operating behavior in response to the services offered. Once that determination is made for all or a sample of participants, it is then necessary to characterize those behavior changes in sufficient detail to estimate associated energy savings. The energy impacts of changes in behavior that result in investment in building retrofits or renewable energy installations can be assessed using the methods discussed under those mechanisms. The energy impacts of some kinds of changes in energy management and facility operation and maintenance practices can also be assessed using a variety of engineering and verification methods. One key aspect of the assessment of evaluability will be a set of preliminary reviews of program records and interviews with principals to determine whether it will be possible to assess the level of measure implementation activity among participants. If that is not possible, then it may be necessary to substitute another PA into this group sample.

The results of verification of changes in behavior and estimates of energy savings for sample participants may be expanded to the program as a whole using a variety of methods associated with simple random samples, stratified samples, or ratio estimators, depending on the nature of the available tracking system data and the kinds of data that can be gathered from participants.

Covered Broad Programmatic Activity Categories and Subcategories. Building Retrofits: Generalized Workshops and Demonstrations; Building Retrofits: Targeted Training and/or Certification; Building Retrofits: Technical Assistance to Building Owners; Loans, Grants, and Incentives: Technical Assistance to Building Owners; Technical Assistance to Building Owners.

4. Improved new construction methods and building system specifications. This energy savings mechanism involves the incorporation into new construction projects of design approaches or pieces of energy end-use equipment that are more energy-

efficient than current standard practice (that is, the non-SEP induced baseline)¹⁷. The principal driver for accelerating changes in building codes in the SEP portfolio (for the ARRA period) is the requirement in the Funding Opportunity Notice that the governor of each state sign an Assurance that the State, or the applicable units of local government that have authority to adopt building codes, will implement:

- A residential building energy code (or codes) that meets or exceeds the most recent International Energy Conservation Code, or achieves equivalent or greater energy savings.
- A commercial building energy code (or codes) throughout the State that meets or exceeds the ANSI/ASHRAE/IESNA Standard 90.1–2007, or achieves equivalent or greater energy savings.
- A plan to achieve 90 percent compliance with the above energy codes within eight years. This plan will include active training and enforcement programs and annual measurement of the rate of compliance.

As of January 2011, 25 states had already have adopted commercial energy codes that meet this standard; and 17 moved to complying residential codes; a few states have already designed or launched effective compliance plans .¹⁸

Many of the building code related PAs reviewed in the PY2008 and ARRA databases focus on training local building code officials in methods to improve compliance.

Generally, the steps in estimating energy savings for these kinds of programs include the following:

- a. Identify the population of new construction projects that may be affected by the changes in code or its enforcement. These might include all new construction and major rehabilitation projects subject to a building code, construction projects designed by architects participating in a training program and so on.
- b. Characterize pre-SEP baseline construction practices for the population of relevant projects. Where appropriate, identify elements of the current code that were changed

¹⁷ We note that, in some cases, past SEP activities may have influenced the current baseline. This would be the case in states where SEP-funded building code upgrade projects. We will note such cases where they occur and make appropriate adjustments to savings as part of the attribution analysis for the PAs in which this situation applies. See Section 4.

¹⁸ Building Code Assistance Project, <http://bcap-ocean.org/code-status-commercial>.

due to past SEP activities so that the gross savings associated with those changes can be evaluated.

- c. Estimate unit savings associated with adopting design practices and/or equipment specifications promoted by the program. Typically, building simulation models or engineering methods are used to estimate the difference in energy consumption between a building or building component designed and built according to pre-SEP baseline practices versus one designed or built according to the standards promoted by the program.
- d. Estimate the number of units actually affected by the program. The approach to this step will vary depending on the program mode. For programs that result in code changes, past evaluations have attempted to project the baseline pace of adoption of changes embodied in the code – that is the hypothetical annual share of projects constructed with those features in the absence of the code changes. Evaluators typically use some combination of in-depth market actor interviews, expert judging, and assessment of analogous programs in the secondary literature to build the estimate. This baseline is then compared to data on actual patterns of construction and code compliance to estimate the number of units affected by the changes promoted by the program. Similar approaches are used to assess the effect of programs aimed at enhancing code enforcement or promoting improved standards for common types of energy equipment.

Covered Broad Programmatic Activity Categories and Subcategories. Building Codes and Standards: all subcategories

Figure 11 summarizes the grouping of BPACs by energy savings mechanism and energy savings estimation methods and rigor levels, along with the preliminary allocation of sample BPACs to these Categories. In the sections that follow we present concise plans for the evaluation of programmatic activities in each BPAC grouping shown in Figure 11.

**Figure 11. Summary of Gross Energy Savings Estimation
and Attribution Approaches by Broad Program Activity Category**

BPAC/Subcategory	n	Rigor Level	Gross Savings Estimation Approach
PY 2008 Sample			
Building Retrofits: Non-res.	2	H	Building Retrofit
Building Retrofits: Res.	2	H	Building Retrofit
Building Retrofits: with General Workshops	5	MH	Building Retrofit/Information
Building Retrofits: with Technical Assistance to Owners	6	MH	Building Retrofit/Information
Loans, Grants, and Incentives: Non-Residential	4	H	Building Retrofit
Loans, Grants, and Incentives: Residential	2	H	Building Retrofit
Loans, Grants, and Incentives: with Technical Assistance to Owners	3	MH	Building Retrofit/Information
Loans, Grants, and Incentives: Alternative Fuels, Ride Share and Traffic Optimization	5	MH	Information
Technical Assistance to Building Owners	3	MH	Information
Renewable Energy Market Development: with General Workshops	6	MH	Renewable Projects/Information
Clean Energy Policy Support	8	MH	Renewable Projects
Building Codes & Standards: Code Development Support	1	MH	Codes & Standards
Building Codes & Standards: General Workshops	3	MH	Codes & Standards
Building Codes & Standards: Targeted Training	2	MH	Codes & Standards
Building Codes & Standards: Technical Assistance to Owners	1	MH	Codes & Standards
Total	53		
ARRA Sample			
Building Retrofits - Non Res	6	H	Building Retrofit
LGI - Non-Res	5	H	Building Retrofit
Building Retrofits - Res	2	H	Building Retrofit
LGI - Res	1	H	Building Retrofit
LGI - Renewable Energy Market Development - Manufacturing	2	MH	Renewable Projects
LGI - Renewable Energy Market Development - Projects	4	MH	Renewable Projects
Renewable Energy Market Development - Projects	4	MH	Renewable Projects
Renewable Energy Market Development - Manufacturing	1	MH	Renewable Projects
Building Codes & Standards: Code Development Support	4	MH	Codes & Standards
Total	29	MH	

4.2 Evaluation Plans: Building Retrofit and Equipment Replacement

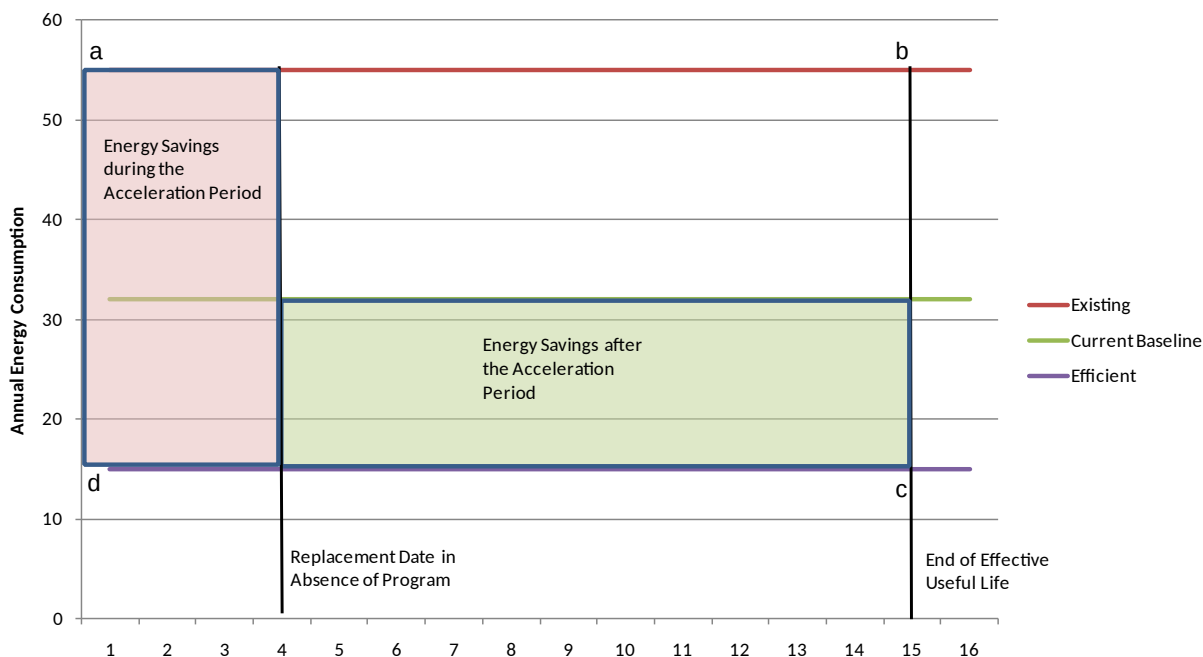
4.2.1 Introduction

Operating definition of energy savings from building retrofit and equipment replacement projects. Based on the review of SEP PA documentation carried out to date and information gained from work on state-level SEP PA evaluations, we have learned that many of the projects supported by the PAs in the Building Retrofit and Equipment Replacement group are retrofit projects. That is, they involve the early replacement of functioning equipment and building systems with efficient models that generally exceed the efficiency new equipment and systems that are considered to be “pre-SEP standard” in the relevant market.

As discussed in Section 1.2, lifetime energy savings is one of the key evaluation metrics for this evaluation. In order to estimate savings from a retrofit project fairly and accurately, it is necessary to determine or to provide clear and reasonable assumptions regarding how long the facility owner would have kept the pre-existing equipment in place in the absence of program assistance to replace it. The example depicted in Figure 12 illustrates the importance of this methodological issue. The solid horizontal lines show the annual energy consumption for a large, durable piece of equipment, such as a chiller, at three levels of efficiency: the equipment in place, the current standard or baseline efficiency for new equipment, and the most efficient equipment available. Assume the program participant installs a new chiller with the highest available efficiency, and that the program induced him to do four years before he would have in the absence of the program. We refer to the period between the program-induced improvements and the (hypothetical) date when they would otherwise have occurred as the “acceleration period.”

During the acceleration period, energy savings would be represented by the shaded area labeled “Energy Savings during the Acceleration Period”. After year four, the relevant efficiency improvement is represented by the distance between the “Pre-SEP Baseline” and “Efficient” annual consumption levels. So, from year four to the end of the equipment’s useful life, the total savings are represented by the shaded area labeled “Energy Savings after the Acceleration Period.” If we had simply projected the savings during the acceleration period to the entire useful life, lifetime energy savings would be much greater, as represented by the rectangle bounded by points a, b, c, and d.

Figure 12. Representation of Energy Savings from Retrofit



In assessing the length of the acceleration and post-installation periods for individual projects or groups of projects, the KEMA team will take the following into consideration.

- Studies of persistence of measures in the field undertaken for public benefits program sponsors.
- Databases of measurement performance such as California's *Database of Energy Efficiency Resources* (DEER) and Technical Resource Manuals that have been developed for other program sponsors.
- Knowledge of the facility management and investment practices of key owner segments. For example, in our own practice we often find that government agencies, operating under budget constraints, retain major heating, mechanical, lighting, and control systems in place well beyond their rated useful lives. Conversely, in retail and office space, lighting systems are replaced frequently with changes in occupancy and mechanical system adjusted to accommodate occupancy needs.

For high-rigor and medium-high-rigor studies, we will gather information directly from program participants to assess the extent to which program assistance accelerated their replacement of

the equipment in question. For high-rigor studies – especially for PAs that include only a small number of large project, this approach may be supplemented by interviews with other decision makers within the participant organization or with vendors who have insight into the owners' baseline practices and the circumstances of the project. For medium-low-rigor studies, we will develop a matrix of assumptions concerning the acceleration period for combinations of measure types and end-user market segments based on the guidance from the information discussed above. This matrix will also serve to fill in “missing values” for sample sites in the higher rigor studies.

Tools for standardization and quality control. According to the sample plan developed in Section 3, the KEMA team will undertake evaluations of 46 separate PAs in the Building Retrofits and Equipment Replacement group. Most of the 25 PAs in the Information Program category (including Technical Assistance) also ultimately target energy reductions through building retrofits and equipment replacement. In order to ensure consistency of methods across these multiple evaluations, transparency of procedures, replicability of results, and an auditable trail for quality control, KEMA is proposing to develop the following database and spreadsheet tools:

- **Savings Calculation Tool.** For all evaluations of PAs that support building retrofit and replacement measures, it will be necessary to develop engineering-based estimates of savings for all or a large sample of participants. For high- and medium-high-rigor evaluations, these engineering estimates will serve as the *ex ante* estimates to which verified savings from a sample of participating sites are compared. For the medium-high-rigor studies, the algorithms used to produce the *ex ante* estimates will be used again with the results of verification data from the sample to estimate verified savings for the sample. For the medium-low-rigor studies, the initial engineering analysis will constitute the full extent of the energy savings assessment.

The PAs in the Building Retrofit and Equipment Replacement and Information Program groups support a broad range of measures in the full spectrum of residential and non-residential end-uses. Moreover, they operate in a wide variety of climate zones and in states characterized by large variations baseline efficiency, as shaped by levels of code adoption and customary building practice. Finally, we know from preliminary work that tracking databases for SEP programs vary greatly in terms of content and quality of data. For example, some databases contain information on square footage of the space in which supported projects are carried out but no other measures of scale, such as counts of units installed. Others contain information on project cost, but no other

measures of size. Our engineering calculations will need to make best use of available local data while making the most reasonable use of assumptions given the nature of the local program, measures, and operating environment. Clearly, the KEMA team will need procedures and tools to manage this diversity while maintaining as much consistency and transparency as possible.

To meet these needs, the KEMA team will develop a Savings Calculation Tool (SCT) for ORNL/DOE. The SCT will be developed in Microsoft Access or Excel and a separate copy populated with local data for each PA evaluation. We anticipate that it will consist of the following components.

- o Savings Algorithm Library. This portion of the tool will contain savings calculations algorithms for the full range of common energy savings measures in the building segments targeted by the PAs. For weather-sensitive measures such as HVAC improvements, the algorithms will include formulae and procedures for taking local weather conditions into account, including specification of the local weather data required. These algorithms will be based on similar work contained in Technical Resource Manuals, as well as the KEMA team's own engineering experience. The sources of all algorithms will be fully documented in this portion of the tool.
- o Input parameter assumption library. This portion of the tool will contain input parameter assumptions used in the algorithms. For some, these will be engineering constants, such as the conversion of motor horsepower to kW or efficiency curves used to estimate savings from VFDs. Others, such as coincidence factors, hours of use for lighting, and heating and cooling degree days will need to be localized to regions, states, or climate zones as appropriate. Finally, this library will contain the "acceleration period" matrices discussed above.
- o Input parameter estimates. This portion of the tool will contain the input parameter estimates actually used in the evaluation of a given PA or project within that PA. These will either be estimated through verification activities for the PA or drawn from the input parameter assumption library.
- o Tracking database file. The tracking database will be copied, moved, or data entered into flat file in the SCT for use in developing *ex ante* estimates of savings at the individual project and PA level of aggregation.

- o Ex ante savings file. This portion of the tool will contain the results of the *ex ante* savings calculations at the lowest level of aggregation supported by the input data. From these results, we should be able to calculate statistics such as savings per project or per unit of various measures that can be used to test the plausibility of estimates and to assess the accuracy of the input data.
- o Verification data file. This portion of the tool will contain the cleaned raw data from the data collection on the verification sample, whether that was done by telephone, on-site, or through some combination of the two.
- o Verified savings file. This file will contain the results of the estimations of verified savings for each sample site. To the extent possible the calculations of verified savings will be stored with the individual site records on this file. For instances in which the calculations are too complex or customized, this file will contain references to work papers and free-standing spreadsheet files.
- o Ratio estimation and sample expansion file. Where ratio estimation is used, this sheet will contain the output of calculations which KEMA generally implements in a statistical package such as SAS. This sheet will also contain the calculations by which the sample data are expanded to the population.
- o Energy savings summary file. This sheet will contain the principal results of the savings analysis, including average annual energy savings, lifetime energy savings, and average peak demand reductions. This sheet may also contain areas for calculations that are driven by energy savings estimates, such as energy cost savings and emission reductions.
- o Cost benefit inputs file. This file will contain the inputs needed for cost benefit analysis and other economic characterizations of the program, including program expenditures, developed in consultation with ORNL/DOE.

KEMA plans to develop this tool first for energy efficiency measures typical of retrofits and equipment replacements in non-commercial buildings, a category of PA that accounts for a significant level of activity. On the basis of experience in building, testing, and using the tool for non-residential retrofit measures, the KEMA team will make a determination as to whether and how to develop similar tools for residential retrofit and replacement measures and for customer-sited renewable energy projects.

Once the tool is created, the project manager of each individual PA evaluation will be responsible for populating it. We will store the current versions of each tool on a central server where senior evaluation managers can have access to them for quality control checks and to verify progress. The populated tool for each PA in which it is used will be submitted as part of the draft Final Report package.

- **Sampling Tool.** KEMA and its subcontractors on this project have developed spreadsheet tools to automate verification sample selection based on accepted strategies such as model-based sampling and so on. Implementation of these methods requires the application of considerable judgment in selecting input parameters to the various formulae that set stratum boundaries, sample allocation, and sample size. These include assessments of the variability in verified savings (versus *ex ante* estimates) for different classes of measures, the costs of different forms of data collection, and implementation of qualitative decisions in sample design, such as mandatory coverage for projects involving certain technologies or located in certain regions.

To support consistency and transparency, KEMA will develop a database of input assumptions for the most frequently-used sample design formulae. These assumptions will include error ratios and coefficients of variation for common types of measures, based to the extent possible on results of recent evaluations. The tool will also contain estimates of costs for various kinds of telephone and on-site data collection and decision rules for implementing various qualitative sampling requirements. The Tracking System and Ex ante Savings files from the Savings Calculation Tool will serve as inputs to the Sampling Tool.

4.2.2 Energy Impacts Assessment Approach

To expedite the description of our approach to assessing the energy impact of PAs in the Building Retrofit and Equipment Replacement grouping, we have consolidated the narrative into its basic steps: assessment of evaluability, processing of the tracking system data, verification sample selection, measurement and verification data collection and analysis, and expansion of sample results to the population of participants. Where there are significant variations in approach for different levels related to rigor levels or specific program type, these are noted and described.

Assessment of Evaluability

Objectives. The objectives of this task are to determine whether it will be possible to evaluate the sampled PA at the rigor level specified by DOE and the sampling process as quickly and with as little expenditure of evaluation resources as possible.

Activities. The Lead Evaluator for the PA will be responsible for collecting information on the criteria listed below and for submitting to the KEMA Project Manager an evaluability assessment within two weeks of initiation of the PA study. The criteria to be applied in assessing evaluability of sampled PAs in this group will include the following:

- Match of actual program operations to BPAC definition. Key program characteristics that distinguish programmatic activities in this BPAC from others in the group include:
 - o Provides support primarily for building retrofit and equipment replacement projects.
 - o The program does not exclusively support measures addressed by other BPACs such as building audits.
- Progress in PA implementation. In order to be considered for evaluation, the programmatic activity needs to have reached the following implementation milestones:
 - o The organization responsible for administering the PA has been identified and, if other than the State Energy Office, has entered into a contract to administer the PA.
 - o Program participant and measure eligibility guidelines and application procedures have been put in place.
 - o For PAs where specific subrecipients were not specified in the funding application, the program has solicited information from eligible subrecipients.
 - o The program is currently active, and is not at risk of cancellation or movement of significant funding to a different BPAC
 - o Program recordkeeping and staff historical knowledge is sufficient to conduct the evaluation
- Progress in project implementation. Determination as to whether a PA in the Building Retrofit BPAC will be included in the evaluation sample will need to be made by July 2011 in order maintain the overall SEP evaluation schedule. By that time, the sampled PA will need to have achieved the following milestones:

- o Received and approved applications, and completed contract agreements for loans or grants (or other applicable incentives) from eligible participants for eligible projects that would commit at least 60 percent of the total funds allocated for incentives.
- o At least 60 percent of the projects that must complete a NEPA approval process have achieved the approval or CX status.
- o Disbursed incentive funding to at least 10 completed projects or, if the projects are very large relative to total project funding, projects that account for 20 percent of total incentive budgets.

Quality and availability of program records. At a minimum, evaluation will require a complete list of participants with contact information, as well as some indicator of the kinds of services and/or incentives received. Immediately upon initiation of the PA level study, the lead evaluator will make an assessment as to whether:

- All or nearly all projects supported by the PA are included in the tracking system, whether that is electronic or paper.
- Whether it will be possible within the budget allocation for the study's rigor level to develop *ex ante* savings estimates given the current state program tracking information and information learned from dialogue with the SEP program manager. Such information would include types of measures installed, end-uses addressed, quantity, efficiency rating, and installed capacity of equipment installed, project costs, and savings estimates developed by other organizations. Where records are not adequate to support the evaluation, the Evaluation Team will work with the program sponsor to upgrade the records as part of the Tracking System Analysis task, for example by review of paper files and other project documentation. If data of sufficient quality to support the evaluation cannot be developed by September 2011, the PA may be dropped from the evaluation sample.

Deliverables. The deliverable for this task will be a memorandum summarizing the Lead Evaluator's findings in regard to the criteria listed above and a recommendation regarding the retention of the PA in the evaluation sample.

Tracking System Analysis and File Review

Objectives. The key objectives of the tracking system analysis and file review task are to:

- Develop *ex ante* savings estimates for all projects included in tracking system.
- Compile and validate other project-level information that will be needed various parts of the survey. This information will include expenditures of SEP funds, participation of other publically funded programs, amounts of matching funding, extent of participant contribution, and contact information for participants and other project principals.

Activities. The principal activities for this task will be as follows:

- **Review tracking system data for completeness and quality.** The first step in the process is to review the entire database to ensure that fields are properly completed to the extent required by the Savings Calculation Tool (SCT) to develop consistent *ex ante* estimates of savings for each project supported by the PA. For larger and more complex projects, review of a sample of project files may be required to ensure that *ex ante* estimates are reasonable. As discussed, we anticipate needing to supplement tracking system data with information gained from paper files and questioning of program staff in many cases.
- **Compile data on local conditions that will be needed to populate the Savings Calculation Tool.** These will include local weather records, verification of current versions of state building codes, and utility cost information.
- **Complete and validate calculations of *ex ante* savings for all projects in the tracking system.** These calculations will be carried out in the Savings Calculation Tool. Once they are completed, members of the PA evaluation team with technical knowledge of the measures involved will review the savings estimates for plausibility. If anomalies are identified, such as savings per unit or square foot of space served that are much larger or smaller than expected, the evaluation staff will review all inputs to the calculations and make adjustments as needed. All adjustments to tracking system inputs will be noted on the *Ex Ante* Savings file.

Deliverables. The deliverables for this task will include the Tracking Database and *Ex Ante* Savings files of the Savings Calculation Tool for the sample PA. The Lead Evaluator will submit these files to the Project Manager along with a memorandum of data quality issues that were encountered in the development of those files and the steps that were taken to address those issues.

Sample Development

Objectives. The objective of this task will be to develop a sample of projects supported by the PA for verification, either by telephone or on-site inspection and measurement.

Activities. The evaluation team will follow standard sampling procedures laid out in guidelines such as California's *Evaluation Framework*¹⁹ for designing stratified samples to support ratio estimation. The evaluation objectives drive the sample design, which will vary depending upon the relative importance of accuracy in the kW versus kWh estimate, levels of aggregation for which results of a given accuracy are sought, and interest in refining savings parameters for particular measures. We use the distribution of projects by tracking system savings and measure type as the basic guides to stratification. If the program addresses a large geographic area and has a sufficiently large population of participants, we may deploy cluster sampling to reduce costs and boost the sample size.

- **Choice of sampling unit.** Generally, we will attempt to match the sampling unit to the purchase decision-making unit in order to capture and make best use of information on attribution of program influence on the quantity of measures, timing, and efficiency levels of equipment installed in direct relationship to the savings estimate. However, this is not always possible due to logistical, schedule, and tracking system problems. We have developed a variety of methods to deal with this problem. For example, we often assess attribution at the program level through large sample surveys of participants, surveys of vendors, sales and shipment data analysis, or combinations of the above.
- **Sample size.** In stratified ratio estimation, sample size is determined by a formula that is driven by the desired level of statistical precision and the underlying variability in the relationship between measured energy savings and tracking system estimates at the sample sites. The statistic that summarizes this variability for a population of projects is known as the error ratio. The error ratio cannot be calculated *a priori* based on a statistical formula related only to sample size. Rather, like the coefficient of variation for a mean, it can only be forecasted on the basis of prior experience in conducting ratio analysis of energy savings from particular kinds of measures and programs. The Evaluation Team has extensive experience in applying this kind of analysis to all of the types of measures and delivery mechanisms encompassed by SEP PAs.

Figure 13 displays the sample sizes required to develop confidence intervals of differing sizes around ratio estimators at the 90 percent confidence level for programs in which there are 200 and 100 participants, which supports effective use of the Finite Population

¹⁹ TecMarket Works, *The California Evaluation Framework*. San Francisco: California Public Utilities Commission. 2004. Chapter 13, Sampling.

Correction. Based on our current knowledge of SEP PAs in this group, we believe that most will have 100 or fewer participants and that 200 is an effective maximum that will be exceeded only in a few large states.

To further limit sampling error, we will use a stratified sampling design, with tracking system savings as the stratification variable. Grouping projects by savings tends to reduce the site-to-site variability of the ratio estimate and increases the precision of the overall savings estimate. Other stratification variables may include measure type and facility type.

Based on experience in evaluating similar types of programs, we believe that use of assumed error ratios in the range of 0.5 to 0.8 will yield samples of sufficient size to provide targeted levels of precision.

Figure 13. Sample Sizes Required for 90 Percent Confidence Intervals Around Ratio Estimators

Underlying Variability	Targeted Precision @ 90 Percent Confidence				
	±10.0%	±12.5%	±15.0%	±17.5%	±20.0%
N = 100					
0.40	30	22	16	12	10
0.50	40	30	23	18	15
0.60	49	38	30	24	19
0.70	57	46	37	30	25
0.80	63	53	44	36	30
0.90	69	58	49	42	35
1.00	73	63	55	47	40
N = 200					
0.40	35	25	17	13	10
0.50	51	35	26	20	16
0.60	65	47	35	28	21
0.70	80	60	46	35	28
0.80	93	71	56	44	35
0.90	105	82	65	53	43
1.00	115	93	75	61	51

Deliverables. As discussed above, KEMA will develop a Sampling Tool for ORNL/DOE that will be used to execute the sample selection. The tool will contain formulae for setting strata boundaries, for allocating sample points to the strata, and for implementing random selection of primary and secondary samples. The inputs to this sheet will consist of the *Ex Ante* Savings file from the Savings Calculation Tool, as well as the Lead Evaluator's instructions for any qualitative criteria to be applied in sample selection. The Sample Tool will complete the shell of the Verification Data file, which will serve as the point of departure for contacting participants in the verification sample. The Lead Evaluator will notify the project manager when the sample is selected and submit a short memorandum summarizing the stratification and sample selection methods used.

Measurement and Verification Data Collection and Analysis

Objectives. The objectives of this task are to develop verified estimates of energy savings for sample projects supported by the PA under evaluation. This step applies only to high- and medium-high-rigor evaluations.

Activities: High-Rigor Studies. Measurement and verification procedures for individual projects will vary depending on the type of measures installed, the percentage of total program savings represented by the site or its stratum, and the level of rigor required. For high-rigor studies, measurement and verification of savings will be accomplished through a combination of on-site inspections, which may include short-term monitoring of key equipment performance factors, with telephone verification interviews. Telephone verification interviews with representatives of sample facilities will validate or update information on the type, quantity, and capacity of equipment measures installed with program support. On-site inspections will include visual verification of measures installed, collection of information to validate baseline assumptions, and, if appropriate and feasible, measurement of impact parameters such as hours of operation or part load using short-term monitoring techniques.

Figure 14 shows the range of activities that will be undertaken for high rigor evaluations of PAs in this group.

Figure 14. Measurement and Verification for Building Retrofit & Equipment Replacement Group: High Rigor

M&V Approach	Data Requirements	Examples
Verification	Verification of installation by interview Project File review for documentation of installation	Small projects and/or projects with simple applications where standard assumptions for operating conditions are appropriate
Verification and engineering savings review	Verification of installation by interview Project file review for documentation of installation Collection of operating conditions and schedule Review of energy savings calculations	Projects where site specific information is available
On-site Installation Verification and Engineering Savings Review	Site Verification of operating conditions and schedule Nameplate data	Sample of largest projects
Metering and measurements	Measurements of key operating parameters for large systems, such as hours of operation, load, supply and return temperatures	Sample of largest projects

M&V Approach	Data Requirements	Examples
	Measurement could be made with already-installed energy management systems	

On-site data collection is expensive. As a rule of thumb, we allocate 24 hours of professional labor and \$400 in direct expenses for travel and equipment rental for site data collection and preparation of verified savings estimates for straightforward projects, such as simple lighting or HVAC equipment replacements in commercial facilities. More complex projects that include a variety of measures or one large measure with many components, such as a chiller system generally require 40 hours of professional labor to analyze, as well as \$600 - \$1,000 in direct expenses. Thus, for programs in the high-rigor category with significant levels of participation, we will probably not be able to conduct on-site inspections and analysis of all projects in the verification sample.

The mix of telephone and on-site verification in the evaluation of a given PA will depend on a number of factors, including:

- Complexity of the sample projects.
- The sample size required for the desired level of precision.
- The distribution of total program *ex ante* savings among individual projects and across technology types.
- The budget requirements of other PA evaluations.

For projects that feature common types of measures such as replacement of lighting fixtures or electric motors, generic site protocols will be used. For sites that feature more complicated projects, KEMA engineers will develop a custom data collection and analysis plan (site plan). These plans take into account the level of rigor specified for the study, number and types of measures installed, the sensitivity of energy savings to variable conditions, including: weather, occupancy, volume of production or facility utilization, customer ability to vary energy service levels, the share of total program energy and demand savings accounted for by the site, and the budget for site data collection and analysis tasks. Development of such detailed site plans usually requires review of materials beyond those contained in the tracking system, including application file materials and direct interviews with facility representatives. Key issues in the process of site-level project verification include the following.

- Determining correct baseline characterization.** In addition to the issues mentioned above, correct baseline specification requires a clear statement of the pre-SEP baseline concept and the means to operationalize that concept. For example, pre-SEP baseline energy use will vary depending on whether equipment replaced had failed or had remaining effective useful life and include estimates of how past SEP initiatives has influenced that baseline. Site protocols must provide for close questioning of participants to ascertain the status of the replaced equipment. See the discussion of this issue in Section 4.1.
- Normalizing results to post-retrofit service levels.** We have found that large energy efficiency projects in both commercial and industrial facilities are often associated with significant changes in production operations or occupancy levels at the site. Thus, the baseline usage must be adjusted to reflect post-retrofit conditions, using production records, customer interviews, energy management system logs, or other data sources.
- Annualization of results.** For a host of reasons, it is generally not feasible to capture directly the effects of variations in weather, occupancy, and production on energy and demand through metering for all or even most of a year. We will use a number of different methods to annualize results from two to six weeks of monitoring. For relatively simple measures, such as lighting or premium efficiency motors, we generally query occupants concerning monthly changes in occupancy schedule and use patterns. Where facility energy management systems monitor hours of use or demand for one or more components of a built-up system, such as a chiller, we will regress measured energy use for the system against that component and weather conditions for the monitoring period, and use the results to model annual energy use and peak demand with a full year of observations from the energy management system. In other instances, we will use site observations and measurements to calibrate hourly building models, such as DOE2.

Activities: Medium Rigor Studies. The M&V methods for the medium high rigor evaluation will also be related to the project type, complexity and percentage of total savings it represents, as shown in Figure 15 below.

Figure 15. Verification for Building Retrofit & Equipment Replacement Group: High Rigor

M&V Approach	Data Requirements	Examples
Verification	Verification of installation by interview	Most projects

M&V Approach	Data Requirements	Examples
	Project File review for documentation of installation	
Verification and engineering savings review	Verification of installation by interview Project file review for documentation of installation Collection of operating conditions and schedule Review and verification of energy savings calculations	Sample of largest projects

Calculation of site-level savings. To calculate the savings, the engineer or analyst will complete these tasks:

- Determine the appropriate baseline conditions.
 - o Normal pre-SEP baseline: The energy savings for retrofits is based on either the pre-existing conditions or on a minimally code compliant replacement that have not been influenced by SEP.
 - o Dual baseline: In the case where the existing equipment was not ready for replacement but was replaced to improve energy efficiency, the remaining useful life of the equipment is considered. The first baseline, the early replacement baseline, uses the energy consumption of the preexisting equipment for the remaining useful life. The second baseline, the normal replacement pre-SEP baseline, applies after the remaining useful life of the equipment until the estimated end of the measure life.
- Perform engineering calculations to determine the gross savings achieved. The gross site savings will be calculated by taking the difference between energy usage for the measure-treated usage and the appropriate pre-SEP baseline. The engineer combines data from the following sources to estimate savings: participant survey interviews, including hours of operation, seasonal patterns of use, control schemes; equipment specifications and invoices; on-site observations; engineering best practices and reference data.

Deliverables. The deliverables for this task will be as follows:

- Verification data file populated with data collected on-site and/or through telephone interviews for each sample project.

- Verified savings file populated with the verified savings estimate for each sample site. This file will also contain references to algorithms and assumptions used from the libraries included in the Savings Calculation Tool, and to external spreadsheets that contain the savings calculations for more complex measures.
- Work papers consisting of savings calculation spreadsheets and scans of paper records, such as manufacturers' cut sheets used in developing savings estimates for complex measures.

Expansion of sample savings estimates: high- and medium-high-rigor studies

Objectives. The objectives of this task are to expand the findings of verified savings from the verification sample to the population of projects supported by the PA.

Activities. The site data analyses yield a set of savings estimates that have been adjusted to reflect findings concerning the actual quantity, efficiency features, operating environment, and operating patterns of the program measures installed in a representative sample of sites. We will use ratio estimation techniques to process these adjusted estimates of savings, along with the tracking system estimates of savings for the sample sites into an estimate of adjusted gross savings for the program as a whole. The ratio estimation will leverage the statistical sample design described earlier and will result in a quantification of program savings with measures of statistical precision and confidence intervals.

The calculation of the adjustment factors for preliminary savings estimates uses appropriate weights corresponding to the sampling rate within each stratum. The two primary adjustment factors are the installation rate and the engineering verification factor. Each of these is calculated as a ratio estimator over the sample of interest. The formulas for these factors are given below.

- G_{Tj} = tracking estimate of gross savings for project j
- G_{Ij} = tracking estimate of gross savings for project j , adjusted for non-installation
- G_{Vj} = verified gross savings for project j based on engineering review.
- A denotes the sample.

Installation Rate

The installation rate R_j is calculated as

$$R_i = \frac{\sum_{j \in A} G_{ij}}{\sum_{j \in A} G_{Tj}}$$

Engineering Verification Factor

The engineering verification factor R_v is calculated as

$$R_v = \frac{\sum_{j \in A} G_{vj}}{\sum_{j \in A} G_{ij}}$$

Deliverables. The deliverables for this task will be the following:

- Ratio estimation and sample expansion file populated with the results of the sample expansion calculations, which will include total energy savings for the PA as a whole, with measures of precision, such as the error ratio, and confidence intervals at confidence levels prescribed by the overall evaluation sample plan. This sheet may also contain estimates of average savings per project or per unit of certain kinds of uniform measures installed. It may also contain findings regarding average values of input parameters such as hours of use or full load hours. These findings may be used to help refine savings parameters used in the Savings Calculator Tool for assessing lower rigor programs.
- Energy savings summary file. This sheet will contain the principal results of the savings analysis, including average annual energy savings, lifetime energy savings, and average peak demand reductions.

4.3 Renewable Energy Market Development Programs

4.3.1 Introduction

As discussed in Section 2, this BPAC encompasses PAs that support the development of individual renewable energy projects—both customer-sited and grid-connected—and PAs that support the development or expanded manufacturing of renewable energy generation equipment. For both of these types of projects, the energy impact assessments will be based on estimation of renewable energy generation and capacity for a sample of installations, and expansion of those estimates to the relevant population of installations using various statistical approaches. The sample PY2008 PAs in this group will be evaluated at a high rigor level,

meaning that at least some of projects installed will be verified on-site. The sample ARRA-funded PAs in this group will be evaluated at the medium-high level, meaning that savings for all supported projects will be verified via remote methods, including telephone interviews with project principals and review of project specifications and energy production records, to the extent those are available.

PAs in the “Manufacturing” BPAC differ from those in the “Projects” in that they do not seek to facilitate direct installation of renewable energy generation equipment. Rather, they achieve energy savings by assisting manufacturers in producing equipment that is more attractive to customers by virtue of lower price, better performance, or features that are more consistent with the planned operating environment. For example, one PA in earlier program years supported a Midwest wind turbine manufacturer in developing more efficient and cheaper turbine blades for small and medium-sized systems, which led in turn to lower customer prices and higher sales in succeeding years. Estimation of savings per turbine or other unit of equipment shipped can be accomplished through various forms of engineering estimates. The more important and difficult task will be to determine the number of units shipped, particularly if production is just getting under way.

Energy savings associated with PAs that provide support for the development of policies that facilitate the development of renewable energy projects will ultimately be generated by such projects. However, DOE has elected to evaluate those PAs at the medium-low-rigor level. Thus, savings for those PAs will be estimated by applying secondary data on average savings per unit of installed capacity to estimates of installed capacity associated with the policy support efforts.

Given the significant differences between the three BPACS in this group in terms of the causal path between program operations and installation of new renewable energy generation equipment, it will be clearer to treat each as a separate program rather than to present them as one program type with a limited number of variations, as we did for the Building Retrofit and Equipment Replacement Group.

4.3.2 Energy Impacts Assessment Approach: Renewable Energy Market Development - Projects

Assessment of Evaluability²⁰

²⁰ To avoid excessive length, we dispense with the Objectives/Activities/Deliverables format for this task.

The criteria to be applied in assessing evaluability of sampled PAs in the Renewable Energy Market Development – Projects BPAC will include the following:

- Match of actual program operations to BPAC definition. Key program characteristics that distinguish programmatic activities in this BPAC from others in the group include:
 - o The funding is allocated to support the installation (not development or manufacturing) of renewable energy technologies.

- Progress in PA implementation. In order to be considered for evaluation, the programmatic activity needs to have reached the following implementation milestones:
 - o The organization responsible for administering the PA has been identified and, if other than the State Energy Office, has entered into a contract to administer the PA.
 - o Program participation and technology eligibility guidelines and application procedures have been put in place.
 - o The program has solicited applications for program support from eligible subrecipients.
 - o The program is currently active, and is not at risk of cancellation or movement of significant funding to a different BPAC
 - o Program recordkeeping and staff historical knowledge is sufficient to conduct the evaluation

- Progress in project implementation. Determination as to whether a PA in the Renewable Energy Market Development -- Projects BPAC will be included in the evaluation sample will need to be made by July 2011 in order maintain the overall SEP evaluation schedule. By that time, the sampled PA will need to have achieved the following milestones:
 - o Received and approved applications, and completed contract agreements for loans or grants (or other applicable incentives) from eligible participants for eligible projects that would commit at least 50 percent of the total funds allocated for incentives.
 - o Approved applications for eligible projects that commit at least 25 percent of the total funds allocated for incentives for the relevant activities.

- o Disbursed incentive funding to at least 10 completed projects or, if the projects are very large relative to total project funding, projects that account for 20 percent of total incentive budgets.

Quality and availability of program records. At a minimum, evaluation will require a complete list of participants with contact information, as well as some indicator of the kinds of services and/or incentives received. The expectation is that PAs can provide the following kinds of information for all projects: equipment installed quantity, installed capacity, project costs, and installation date. Where records are not adequate to support a high- or medium-high-rigor evaluation, the Evaluation Team will work with the program sponsor to upgrade the records by incorporating data from paper records or direct contact with participating facility owners. If adequate tracking system data cannot be developed by September 2011, the PA may be dropped from the evaluation sample.

Tracking System Analysis and File Review

Objectives. The key objectives of the tracking system analysis and file review task are to:

- Develop *ex ante* estimates of renewable energy installed capacity and generation for all projects included in tracking system.
- Compile and validate other project-level information that will be needed for various parts of the survey. This information will include expenditures of SEP funds, participation of other publically funded programs, amounts of matching funding, extent of participant contribution, and contact information for participants and other project principals.

Activities. The principal activities for this task will be largely the same as they are for the Building Retrofit and Equipment Replacement group.

- **Review tracking system data for completeness and quality.** The first step in the process is to review the entire database to ensure that fields are properly completed to the extent required by standard engineering techniques to estimate *ex ante* renewable energy generation for each project supported by the PA. We anticipate needing to supplement tracking system data with information gained from paper files and questioning of program staff in many cases. One key data element that will be needed is the availability of meter information on renewable energy generation for each project. Accurate information on the type of metering equipment installed and the nature of records retained will be needed for planning the verification step.

- **Compile data on local conditions that will be needed to carry out engineering estimates of renewable energy generation at the project level.** These will include local weather records; sun and wind resource statistics; local interconnection requirements, regulations, and tariffs; verification of current versions of state building codes, and utility cost information.
- **Complete and validate calculations of *ex ante* renewable energy generation for all projects in the tracking system.** An alternative to expansion of sample results on the basis of estimated energy generation would be to use capacity installed, if those data are available for all projects.

Deliverables. The deliverables for this task will include cleaned files of the tracking system database and either estimates of renewable energy generation for all projects or validated measures of size for all projects, depending on the sample expansion approach to be used. The Lead Evaluator will submit these files to the Project Manager along with a memorandum of data quality issues that were encountered in the development of those files and the steps that were taken to address those issues.

Sample Development

Objectives. The objective of this task will be to develop a sample of projects supported by the PA for verification, either by telephone or on-site inspection and measurement.

Activities. The sampling approach and activities deployed for evaluations of PAs in this BPAC will be similar to those used in the Building Retrofit and Equipment Replacement group. Some sampling issues particular to this group are as follows.

- **Stratification variables.** The stratification variables for this group will certainly include technology (if the PA supports more than one) and size as measured by installed capacity or *ex ante* generation. Other variables such as market segment of the owner (residential versus non-residential) or elements of the operating environment may also influence generation or variability of generation.
- **Measure of size.** As discussed earlier, it may be easier and more consistent to use installed capacity as the measure size, rather than estimates of energy generation, which will in any case be closely related to installed capacity.

Deliverables. We anticipate that we will be able to use the Sampling Tool described above to carry out selection of renewable energy projects for verification. See the discussion of sampling deliverables for the Building Retrofit and Equipment Replacement group above for a description of the output of that tool. The Lead Evaluator will notify the project manager when the sample is selected and submit a short memorandum summarizing the stratification and sample selection methods used.

Verification Data Collection and Analysis

Objectives. The objective of this task is to develop verified estimates of renewable energy generated for sample projects supported by the PA under evaluation. This step applies only to high and medium-high rigor evaluations.

Activities: High Rigor Studies. Renewable energy generation for a given project can be estimated in two ways:

- Metering and monitoring system – Some systems, especially large ones, have metering and monitoring systems which record the system’s production data. The data may be housed locally or at a remote site with a meter data monitoring provider. The KEMA team will ascertain the kind of meter data available for each sample site as part of the tracking system review.
- Engineering estimates using on-site data – System production can be estimated by measuring a site’s resource availability, its system design, and the equipment used.

Verification procedures for individual projects will vary depending on the type of measures installed, the percentage of total program savings represented by the site or its stratum, and the level of rigor required. The verification protocols, whether on-site or telephone, will aim to validate type and quantity of equipment installed. On-site inspection will be needed to verify other kinds of conditions that affect equipment performance, such as the quality of installation, maintenance of the equipment, and ambient conditions shading. For projects that feature common renewable energy system such as roof-mounted PV and small wind, generic site protocols will be used. For projects that feature less common renewable systems, Evaluation Team engineers will develop a custom data collection and analysis plan.

For high rigor studies, measurement and verification of savings will be accomplished through a combination of on-site inspections and telephone verification interviews. Telephone verification interviews with representatives of sample facilities will validate or update information on the type, quantity, and capacity of equipment measures installed with program support. As

discussed in regard to Building Retrofit PAs, the Lead Evaluator will determine which sampled projects will require on-site validation based on a number of considerations, including:

- The nature of the systems installed;
- The quality of tracking data available on the capacity and other key features of the systems installed;
- The quality of metered data available.

Activities: Medium-High-Rigor Studies. For the medium-rigor studies, verification information will be collected only through remote activities, including file review and interviews with project owners and operators. If the PA supported only a few very large projects and it is not possible to characterize those projects in detail from information such as records of renewable energy generation, the Lead Evaluator may recommend on-site data collection to characterize the physical installation in sufficient detail to support engineering estimates of generation.

Expansion of sample savings estimates to the population of projects

The sample expansion procedures to be used in the evaluation of PAs in this group are the same as those described for the Building Retrofit and Equipment Replacement group.

4.3.3 Energy Impacts Assessment Approach: Renewable Energy Market Development - Manufacturing

Assessment of Evaluability

The criteria to be applied in assessing evaluability of sampled PAs in the Renewable Energy Market Development – Manufacturing BPAC will include the following:

- Match of actual program operations to BPAC definition. Key program characteristics that distinguish programmatic activities in this BPAC from others are that funding is allocated to support the development or manufacturing of renewable energy technologies.
- Progress in PA implementation. In order to be considered for evaluation, the programmatic activity needs to have reached the following implementation milestones:

- o The organization responsible for administering the PA has been identified and, if other than the State Energy Office, has entered into a contract to administer the PA.
 - o Agreements with the organizations receiving support that specify the uses of the funding need to have been signed by all relevant parties.
- Progress in project implementation. Determination as to whether a PA in the Renewable Energy Market Development – Manufacturing BPAC will be included in the evaluation sample will need to be made by June 2011 in order maintain the overall SEP evaluation schedule. By that time, at least one of the organizations (e.g., grantee and/or other supporting organizations) receiving product development and manufacturing support will need to have achieved the following milestones:
 - o Developed the product or product feature for which it received support to the point where it can be sold commercially.
 - o Initiated sales activities in support of the new or improved product.
 - o Obtained at least one order for the new or improved product.

Without this minimum level of commercial experience with the product, it will not be possible to advance credible claims concerning renewable energy generation or to forecast potential acceptance of the products using expert opinion as a guide.

- Quality and availability of program records. The following records will be required to support the evaluation.
 - o Agreements between the state energy office and subrecipients specifying the technologies to be supported, milestone accomplishments, conditions of payment, and so forth.
 - o Contact information for the subrecipients and any other key project principals and consultants.

Successful completion of these PA evaluations will depend on close cooperation from subrecipients in characterizing the progress made using the SEP funds; the effect of the SEP funding on product features, performance, production costs, or asking prices; the markets targeted by the technologies; and contact information from customers who would be willing to be interviewed concerning the effect of the availability of improved or

less expensive on their purchase and installation decisions. If such cooperation is not forthcoming it may be necessary to drop the PA from the evaluation sample.

Development of Preliminary Renewable Energy Generation Estimates

Objectives. At this point, the methods proposed to estimate renewable energy generation for manufacturing PAs depart significantly from those proposed to estimate generation from project-oriented PAs. The objective of this first step in the process is to develop the following for each manufacturer who received support from the sampled PA:

- An algorithm for energy generation per unit of capacity installed that can be calibrated to conditions at the locations in which the new or improved equipment is installed, for example: local wind resources, local solar resources, mix of biomass inputs available.
- A forecast of sales over the period ending 2016 of the products whose development and/or manufacture received support from the program, with assessments of market size, competitors' offerings, trends in capacity cost, and other information that can help to support the forecast. This forecast will be used to quantify the number of units installed and to support an expert judging process to be carried out as part of the attribution assessment.

Activities. The Lead Evaluator and staff familiar with the technology and markets for renewable energy equipment will work closely with principals of the subrecipients and state energy office officials to develop the savings algorithm and the sales forecast. If the subrecipient cannot or will not cooperate in the development of these inputs, the project will be dropped from the sample. We also believe that any savings claims from the project would need to be severely discounted if the principals cannot contribute to developing these inputs, pending ORNL approval.

Deliverables. The deliverables for this task will be the algorithm for energy generation per unit installed and the sales forecast for each product whose development or manufacturing were supported by the sampled PA.

Development of Early-Year Verified Estimates of Renewable Energy Generation

Objectives. The objectives of this task are to develop verified estimates of energy production per unit of capacity for installations of the subject technologies completed or in substantial progress prior to the end of the evaluation period. The data collection for this task will also

support the attribution phase by collecting project owners' or installers' assessments of the importance in their purchase decisions of product features supported by the PA.

Activities. The activities for this task will include the following:

- Identification of purchaser sample. Working with the subrecipients, the KEMA team will identify firms that purchased the renewable energy equipment affected by PA support, whether for installation in their own facilities or for installation in customer facilities.
- Collection of information on installations. Evaluation staff will interview up to five firms in the purchaser sample to ascertain the number of installations completed using the technology under review, and the capacity, location, and institutional setting of typical installations, focusing on attributes that will affect savings achieved.
- Collection of attribution information. As part of the purchaser interviews, evaluation staff will question respondents on the following topics:
 - o Reasons for selection of the equipment versus competitor offerings.
 - o Comparison to competitor offerings in terms price, performance, and features.
 - o Likely course of action if equipment corresponding to the manufacturer's specific offering had not been available.

Deliverables. The deliverables for this task will be verified estimates of renewable energy generation for the sampled purchasers and a record of the results of the attribution interviews.

Development of Renewable Energy Generation Forecasts

Objectives. The objective of this task is to develop long-term estimates of the renewable energy generated by new or improved equipment installed as a result of support provided to manufactures through SEP PAs. If, on the basis of interviews with project principals and state energy officials, the Lead Evaluator concludes that the effect on equipment sales of the development and manufacturing activities undertaken with SEP funding will have run its course by March 2012, then PA level savings can be projected on the basis of sales or projected sales through that period and the estimate of energy generated per unit of capacity installed developed through the early year verification efforts describe above. If, on the other hand, sales of the new or improved product are growing in March 2012, when KEMA needs to conclude field

research, we will need to forecast the number of units sold over a reasonable time horizon. Unless additional specific information is available, for purposes of this plan, we propose using a sales forecast horizon of five years, closing at 2017, beyond which we believe forecasts would not be reliable. Savings for all units sold during this time period will be counted for their full effective useful lives (EULs). See the discussion of attribution of savings for manufacturing oriented programs in Section 5.3 for details of this approach.

4.3.4 Energy Impacts Assessment Approach: Renewable Energy Market Development – Clean Energy Policy Support

Assessment of Evaluability

The criteria to be applied in assessing evaluability of sampled PAs in the Clean Energy Policy Support BPAC will include the following:

- Progress in PA implementation. In order to be considered for evaluation, the programmatic activity needs to have reached the following implementation milestones:
 - o *If implementation of the policy requires regulatory decisions or mandates:* The relevant regulatory body has issued a written decision or order requiring implementation of the policy in question, clearly identifying the actions that need to be taken as well as the parties to be held responsible.
 - o *If implementation of the policy requires legislative action:* The state legislature has passed enabling legislation and provided funding, if necessary, in the current program year.
 - o *If the policy requires ongoing administrative oversight or delivery of public services:* Responsibility for administrative oversight or program delivery has been assigned to a state agency or quasi-public authority, required functions are staffed, and program operation has been underway for at least six months.
- Progress in project implementation. Determination as to whether a PA in the Renewable Energy Market Development – Manufacturing BPAC will be included in the evaluation sample will need to be made by July 2011 in order to maintain the overall SEP evaluation schedule. By that time, at least one of the following milestones in project development needs to have been achieved.

- o *If the policies target the acceleration of investment in relatively small customer-sited clean energy installations:* State energy officials need to be able to identify at least five projects that have been completed or are that are currently under development and that – arguably -- have been facilitated by the policy under evaluation. The Lead Evaluator for the PA will assess this latter claim in light of the features of the project and the stated objectives of the policy prior to recommending whether to proceed with the PA study.
- o *If the policies target the acceleration of investment in large-scale grid-connected projects.* State energy officials need to be able to identify at least one such project that is completed or in development that was facilitated by the adoption of the policy under evaluation.

Estimation of energy impacts

Once the Lead Evaluator has established the evaluability of the sample of Clean Energy Policy Support PA, estimation of energy impacts will proceed in the following steps.

- Characterize the population of projects whose development was facilitated by the policy under evaluation. This initial task will be accomplished through interviews with state energy officials, representatives of renewable energy industry associations, and firms that sell the relevant equipment and project development services. Evaluation team staff will contact a sample of up to 9 project owners to gather information on capacity installed and other operating characteristics. This information will be used to develop estimates of the average capacity installed for a typical project and of energy generated per unit of capacity installed. If such estimates cannot be gathered from project principals, the evaluation team will develop them using engineering-based methods.
- Estimate average annual and lifetime renewable energy generation for identified projects. The evaluation team will use the data on installed capacity and project type to develop estimates of annual and lifetime energy generation for the identified projects. The estimate of renewable energy generation will be adjusted per the results of the attribution analysis, as discussed in Section 5.3.

4.4 Information and Training Programs

4.4.1 Introduction

Assessment of energy savings for PAs in these groups will rely on many of the same devices as discussed under the Building Retrofit and Equipment Replacement group. The main difference is that estimation of energy impacts for Information and Training programs and components of programs evaluated under the Building Retrofit and Equipment Replacement approach may require that we estimate the portion of participants who undertook energy efficiency or renewable energy measures and characterize those measures through contact with a sample of participants. Once that step is accomplished, we will apply methods discussed under the medium-high-rigor portions of the Building Retrofit and Equipment Replacement group to estimate energy savings.

4.4.2 Energy Impacts Assessment Approach

Evaluability Assessment

The assessment of the evaluability of the program efforts will depend on the following:

- The status of program implementation. In order to be considered for evaluation, the Programmatic Activity needs to have reached the following implementation milestones:
 - o The organization responsible for administering the program has been identified and, if other than the State Energy Office, has entered into a contract to administer the program.
 - o Program marketing and outreach materials have been developed and launched.
 - o Program curriculum and other content have been developed, where applicable.
 - o Activities have been available to potential participants for at least the last six months of 2010, e.g. held trainings or seminars, provided events, etc.

Some types of programs will experience long lead times until energy, environmental, and employment impacts occur. If implementation of a selected program is significantly

delayed, then we may reallocate limited evaluation resources to another program from the BPAC.

- The matching of actual program operations to BPAC definition. In this step, the Evaluation Team will confirm that the databases are correct and that the Programmatic Activity falls within WE&T.
- The quality and availability of program records. Generally, evaluation of programmatic activities in the WE&T BPAC will require sufficient data on activities supported by the program administrator. At a minimum, evaluation will require a complete list of programmatic activities, expenditures per activity, contact information for all state level program administrators, contact information for trainers/implementers (where available), materials or curricula, and some indicator of the kinds of activities offered. Beyond that, information on the kinds of projects implemented and expected levels of energy savings will increase the likelihood of completing an evaluation with an acceptable level of rigor. Further, lists of program participants, with contact information will increase the likelihood of identifying program influence to assign attribution with a high degree of confidence. If the records for a sampled program are judged to be inadequate to support an evaluation of the required rigor, the Evaluation Team may request that the program sponsor make necessary changes and additions to record-keeping processes. If those changes cannot be effected within a specified time period, the program may be dropped from the evaluation sample.

Tracking System and Program Records Analysis

Objectives. The key objectives of the tracking system analysis and program records review task are to:

- Identify the content of all training and technical services provided (e.g. course curricula) and establish the linkages to specific groups of market actors, end-uses targeted, and energy efficiency or renewable energy measures promoted.
- Compile records of all recipients of information and training services: attendees at workshops, facility owners receiving technical assistance, users of technical information clearing houses and so forth. To the extent possible, associate all participants with standardized descriptors of the services they received.

Activities. The principal activities for this task will be as follows.

- **Review tracking system data for completeness and quality.** The first step in the process is to review all program records to ascertain the availability of the information described above under objectives. At a minimum we will attempt to acquire the following materials for each activity:
 - o Course and activity catalogs
 - o Course and activity descriptions
 - o Course curriculum
 - o Marketing materials
 - o Application forms
 - o Program implementation plans, where available
 - o Quarterly reports
 - o Past evaluation reports
 - o All available activity tracking databases, including workshop volume, workshop instructors or facilitators, and detailed participant data
- **Develop sample frame for participant surveys.** Using this information, we will develop, to the extent possible, a table that is analogous to the Tracking System file in the Savings Calculation Tool discussed above. This file will contain the contact information for individual program participants along with variables describing the services they received and the dates on which they received them. This file will serve as the sample frame for interviews or surveys of program participants.

Deliverables. The deliverables for this task will include a summary of all information services provided by the PAs, including counts of participants or other measures of volume of activity, such as hits on a clearinghouse website where such information is available.

Sample Development

Objectives. The objective of this task will be to develop a sample of program participants for contact via telephone or, if appropriate, e-mail.

Activities. The principal activities for this task will be as follows.

- **Develop the sample design.** The sample of participants will be stratified by variables that are associated with the level of expected energy savings. At a minimum we will stratify by type of service received. If program records contain any information on

individual participants that would support even rough *ex ante* estimates of savings, such as type and square footage of the facilities in which they work, then we may use size as a stratification variable as well. See the discussion of sample structure in the Building Retrofit and Equipment Replacement section for detail on sample stratification and size.

- **Select the sample.** The primary and secondary samples will be selected using standard random sampling procedures.

Deliverables. As discussed above, KEMA will develop a Sampling Tool spreadsheet that will be used to execute the sample selection. The tool will contain formulae for setting strata boundaries, for allocating sample points to the strata, and for implementing random selection of primary and secondary samples. The Sample Tool will complete the shell of the Verification Data file, which will serve as the point of departure for contacting participants in the verification sample. The Lead Evaluator will notify the project manager when the sample is selected and submit a short memorandum summarizing the stratification and sample selection methods used.

Verification Data Collection and Analysis

Objectives. The objectives of this task are to develop verified estimates of energy savings for sample projects supported by the PA under evaluation.

Activities: The principal activities for this task will be as follows.

- **Complete a measure implementation survey with the participant sample.** In order to promote consistency in analysis and reduce overall respondent burden, we plan to use the same survey form to gather information on measure implementation as will be used in the medium-high-rigor studies in the Building Retrofit and Equipment Replacement group. This survey will first verify that the respondent took part in the information and training program under evaluation. Once the correct respondent is identified, the survey will proceed to questions that characterize which measures related to that training have been implemented in the time since participation, as well as the respondent's perceptions of the effect of the information and training services received on his or her organization's decision to implement those projects. Finally, the survey will elicit information on the nature of the measures installed, their quantity, and efficiency specifications, as well as relevant details on the facility's physical features and occupancy needed to support engineering estimates of energy savings.

- **Develop engineering estimates of savings.** The evaluation team will use the Savings Calculation Tool in combination with data collected through the measure implementation survey to estimate energy savings for each sampled participant. If the participant has not implemented any measures to which the information and training program was relevant, the savings will be counted as zero.
- **High-rigor studies: on-site verification.** For high-rigor studies, on-site verification of a small subsample of participant sites may be authorized if a participant reports implementing measures that could account for 3 percent or more of total program savings.

Calculation of site-level savings. See the discussion of calculation of site-level savings in the Building Retrofit and Equipment Replacement section.

Deliverables. The deliverables for this task will be as follows.

- Implementation data file populated with data collected on-site and/or through telephone interviews for each sample project.
- Verified savings file populated with the verified savings estimate for each sample site. This file will also contain references to algorithms and assumptions used from the libraries included in the Savings Calculation Tool, and to external spreadsheets that contain the savings calculations for more complex measures.
- Work papers consisting of savings calculation spreadsheets and scans of paper records, such as manufacturers' cut sheets used in developing savings estimates for complex measures.

Expansion of sample savings estimates: high and medium high rigor studies

See the discussion of expansion of sample savings in the Building Retrofit and Equipment Replacement section. Our analysis of the implementation data for these programs will also include detail on the percentage of participants in various programs who went on to implement measures, the characteristics of those measures, and participant perceptions of the effects of the information programs on decisions to implement the measures specifically supported by the program, as well as other more general energy efficiency investments and energy management strategies.

4.5 Codes and Standards Programs

4.5.1 Introduction

The details of the Codes and Standards (C&S) classed programs funded by SEP vary significantly across the States and entities. In PY2008 most of the C&S programs included advocacy for the adoption of energy efficiency codes. A few States also included development of codes specific to the States, usually based on one or more existing model codes. In addition, many programs included code official and builder/developer training on code compliance. In PY2009-2011, participation in the ARRA-funded SEP required States to have adopted State-wide energy efficiency codes as a precondition for ARRA-enhanced SEP participation. In addition, the participants were required to demonstrate 90 percent compliance with the adopted codes. This fundamental change in building energy efficiency code coverage and enforcement driven by the ARRA program is likely the largest single impact of the C&S programs. Thus, we plan to evaluate the effects of code change efforts in a sample of states according to the sampling plan as well as PAs that involve other aspects of code-related work, such as training for code enforcement officials and other enhancements to enforcement systems.

4.5.2 Estimation of Potential Energy Effects

The estimation of the total pool of savings available from strategies that advance the date of new code adoption will be accomplished through the following steps.

Identify the population of new construction projects that potentially may be affected by the program.

For the sampled States, construction activities will need to be determined individually. Many States may maintain adequate construction records while for other States the evaluation team will need to seek out secondary records from which construction activity can be estimated. It is probable that for all States in the sample, major construction activity will be dominated by a few jurisdictions which will simplify sample selection. For each affected State, no more than nine code offices will be visited and their code compliance documentation checked for the entire 2009-2011 period. The compliance confirmation report for the affected States will be reviewed and compared to our findings in the offices. No individual building site inspections will be conducted.

Characterize pre-SEP baseline construction practices for the population of relevant projects.

Development and adoption of building energy codes and standards has been an ongoing process for the last several decades. Some States like California have had SEP influenced energy efficiency codes in place for more than 30 years, but many States had not adopted building energy efficiency codes until the ARRA funding preconditions were instituted. In conducting research on state codes, we will review revisions made in the previous five years and characterize the role played in those revisions by SEP officials or grantees. There is sufficient history and understanding in the industry for experts to characterize with reasonable certainty common practices in each of the States prior to 2009-2011 period.

The pre-SEP baseline construction practices in each affected State will be determined through interviews with a small panel of industry experts familiar with the SEP change efforts and their cause and effect relationships, but also who are familiar with national practices and understanding of local conditions and practices. Information from these sources will be supplemented by research in secondary sources, including residential appliance saturation surveys, construction market studies, census data, and information from the major federal surveys of energy use including the Residential Energy Consumption Survey (RECS) and the Commercial Building Energy Consumption Survey (CBECS).

Estimate unit savings associated with adopting design practices and/or equipment specifications promoted by the program.

The limited information that is expected to be available to characterize the building population affected by the C&S Programs means the evaluation will need to develop a simplified estimating technique. Potentially, the affected code offices will have compliance models (building energy simulations) that could be used. These models could be modified to reflect the probable choices that would have been made prior to the existence of the energy code. The difference would represent the percentage change in energy efficiency induced by the adopted energy efficiency code. There could be simpler approaches used depending on the findings at each individual code official office. The large variation in practices by local code offices will probably result in large uncertainties. Typically, engineering methods are used to estimate the difference in energy consumption between a building or building component designed and built according to baseline practices versus one designed or built according to the standards promoted by the program. The relevant indicator of unit energy consumption will vary depending on the targeted population or technology. For example, it could vary by annual consumption per housing unit

for residential programs, annual consumption per square foot for commercial buildings, or annual consumption per Btu/hour capacity for HVAC equipment. The engineering methods applied to estimate unit energy savings will vary according to rigor level, and include simple parametric calculations based on secondary data and various types of building or system simulation modeling.

Estimate the number of units affected by the program.

The impact of C&S is dependent on the market for new and renovated buildings, which is in turn dependent on the overall business cycle. States vary in the kind of construction activity records they maintain; however, since this is focused on buildings subject to the building code it should be possible to obtain copies of construction and occupancy permits from the local code official offices. That information will provide the basic information on the number of permitted and the number of completed buildings. Not all permitted buildings are constructed and sometimes there are significant delays between permitting and occupancy. The basic data for estimating the number of units affected by the program; in the new construction and major renovation activity, will be estimated at a first order of approximation by analysis of project data maintained by FW Dodge in its *Players* database. We have found the *Players* database to be a suitable resource for studies of this type.

5. Attribution Approaches

This chapter presents the KEMA team's proposed approach to assessing attribution of estimated energy savings to the sample SEP PAs. The introduction to this section states the fundamental research questions underlying the attribution assessment, elaborates on their relation to specifics of SEP objectives, PA offerings, and operating environment, and provides an overview of methods available for addressing the fundamental research questions. The subsequent sections of the chapter summarize our proposed application of the basic framework to evaluation of PAs in the groups established in Section 4. To expedite the presentation, we have consolidated the PA groupings into two sections that share similar methodological approaches and challenges. These are:

- **Group 1: Building Retrofit and Equipment Replacement, Renewable Energy Market Development – Projects, Information and Training Programs.** These programs focus on providing individual market actors with the information, tools, and incentives they may need to accelerate the adoption of targeted energy efficiency and renewable energy measures in specific projects. In assessing attribution for these programs, we will rely heavily on information gathered directly from program participants who are key decision makers in the financial decision, especially in high and medium-high rigor evaluations. These data on participant perceptions of program influence will be supplemented by information from vendors, program managers, and other market observers. We will screen potential interviewees among other market actors to assure that they are at least aware of SEP-supported program activities (if not of the connection of those activities to SEP funding). Without such awareness the market actors would be unlikely to be able to comment on the extent of the influence of SEP-supported activities.
- **Group 2: Renewable Energy Market Development – Manufacturing, Clean Energy Policy Support, Codes and Standards.** These programs address projects as individual transactions. Rather, they attempt to influence large classes of projects by establishing favorable conditions for their implementation by improving the performance and cost-competitiveness of efficient technologies (manufacturing-oriented programs) or by removing barriers and creating incentives through regulatory and policy initiatives. Alternatively, they may oblige whole classes of customers to adopt efficient technologies through their incorporation into building codes and equipment standards. For these types of PAs, the perceptions of individual facility owners will provide little insight into attribution of observed savings. Rather, the attribution analyses for these programs will rely heavily on the collection, compilation, and interpretation of perceptions and opinions

from knowledgeable supply side market participants and market observers, including regulators and code officials. This information will be supplemented by research into secondary sources that trace the development of the relevant markets. These analyses will make extensive use of logic models and other devices discussed below for structuring case study materials.

5.1 Introduction

5.1.1 Fundamental Research Questions

Overview. The KEMA team has identified three fundamental research questions to be addressed in the attribution assessment for each sample PA. These are as follows:

1. ***What would the market actors targeted by the sample PA have done in regard to adopting the PA-supported technology or service in the absence of the program?*** This is the classic question posed in evaluations of all kinds of programs. It provides the framework for assessing the attribution of observed changes in key outcomes to the effects of the program.
2. ***In instances when two or more programs, including the SEP PA, target the same outcomes in the same domain²¹, to what extent are observed outcomes attributable to one program or another?*** This question is particularly important in the case of the evaluation of SEP PAs for a number of reasons. First, in many states, ratepayer funded programs with significantly greater resources targeted some of the same outcomes, particularly in the pre-ARRA period but also in the ARRA period. Second, to leverage its resources, SEP PAs often coordinate explicitly with programs offered by other sponsors which provide additional resources for efficiency and renewable measure adoption. State energy officials believe that their SEP programs have influenced their target markets to an extent far greater than would be suggested by their level of funding. It will be necessary to test formulate and test such hypotheses in the individual PA evaluations.
3. ***To what extent have SEP PAs influenced the allocation and deployment of resources by other program sponsors in the relevant domains?*** A number of studies of SEP activities²² have found that sponsors of ratepayer-funded programs collaborated closely with state energy offices to leverage their own resources, especially

²¹ By “domain” we mean the groups of market actors, regulators, government bodies and other institutions and their network of interactions in which the program operates and that it attempts to influence.

with the influx of ARRA funding. This means that, “in the absence of the program,” the array of resources available to market actors in the PA domain would have been reduced not only by the absence of the SEP PA activities, but by a reduction in the level of resources available from other program sponsors. Thus, it will be necessary to formulate and test hypotheses regarding the influence of SEP PA activities on the programming decisions of other sponsors in the domain. The findings from this analysis may be used to inform research to address the Research Question #1.

The following paragraphs elaborate on these questions within the specific context of SEP activities in PY2008 and the ARRA period.

Apparent program effects on market actors. As the analysis in Section 2 shows, the programmatic activities supported by SEP, both in the pre-ARRA and ARRA periods, are extremely diverse. However, they all have the same basic objective, namely to encourage actors in the markets for energy and related capital goods and services to adopt energy-efficient and renewable energy technologies and practices. Market actors in this case include energy users as well as firms and individuals in the supply chain for energy using equipment, renewable energy generating equipment, and design, installation, and maintenance services. Thus, as in evaluations of most energy efficiency programs, the key question to be addressed in assessing the attribution of estimated energy impacts to SEP programmatic activities is this:

What would the targeted market actors have done in regard to adoption of the supported technology or service in the absence of the program?

Over the past 25 years, evaluators of energy efficiency programs have developed a repertoire of methods to address this question, mostly involving incentive programs operated by utility companies. These methods and their applications are summarized in Section 5.1.2.

Relative influence of other programs active in the sample PA's domain. Most evaluations of energy efficiency programs take into account the potential influence of programs and policies other than the ones under evaluation on the outcomes of interest, such as the change in the pace of adoption of the targeted technology. This is typically accomplished through some type of quasi-experimental research design or by explicitly probing the influence of other programs

²² TecMarket Works. *The State Energy Program: Building Energy Efficiency and Renewable Energy Capacity in the States*. Oak Ridge, TN: Oak Ridge National Laboratory. September 30, 2010. And Goldman, Charles A. et al. *Interactions between Energy Efficiency Programs funded under the Recovery Act and Ratepayer-funded Energy Efficiency Programs (Draft)*. Berkeley CA: Lawrence Berkeley National Laboratory. January, 2011.

and policies in surveys and interviews with a sample of market actors. The SEP approach to program design, the resources available to SEP, and the operating environments of many PAs elevate the importance of this issue beyond its level in most evaluations. Specifically:

Pre-ARRA Period

- **Levels of funding.** In 2008, roughly 18 states had well-established ratepayer funded energy efficiency programs in operation spending tens and in some cases hundreds of millions of dollars per year on outreach and incentives. Another 10-12 states were in the process of deploying new ratepayer-funded programs or reviving programs that had been dormant. In these states, the level of energy efficiency and renewable spending and activity by ratepayer funded programs was far larger than that of SEP related PAs.
- **Programming strategies.** During the pre-ARRA period, state energy offices generally followed a number of strategies to generate the greatest benefits from their limited funding. These strategies included:
 - o Focus on targeted technical support projects to advance changes in regulations that have far-reaching impact on adoption of energy efficiency and renewable energy technologies: codes and standards, renewable portfolio standards, interconnection rules and tariffs, etc.
 - o Target programmatic activities to energy efficiency opportunities that ratepayer programs generally do not address due to their cost-effectiveness frameworks. These include programs that save unregulated fuels such as oil or programs that serve small, hard-to-reach customer segments.
 - o Design and deliver programs that steer market actors into participation in the ratepayer funded projects, whether explicitly or not. For example, many of the PAs in the Workshop, Training, and Education BPAC provided training to commercial facility owners to identify opportunities for improvements to HVAC or control systems, followed by guidance in seeking incentives from ratepayer funded programs to implement those measures.

In conducting attribution analyses of programs of this last type, it will be necessary to assess the following questions:

- o What percentage of training program participants went on to implement relevant projects?
- o What percentage of those who implemented projects sought and obtained support from other programs to do so?

- o What percentage of those who implemented with the support of other programs were aware of those programs prior to participation in the SEP PA?
- o What relative level of importance or influence do SEP PA participants assign to the various programs in their decision to adopt the technologies in question?

It may be reasonable to hypothesize that the SEP PA had an influence on market actors greater than would be indicated by its funding relative to other programs, for example, due to its earlier access to the market actors or its efficacy in overcoming information as opposed to financial barriers. However, in this evaluation such hypotheses will need to be formulated in the specific context of the sample PAs and tested using the established tools of social science research.

ARRA Period

- **Levels of funding.** Even after the massive short-term infusion of ARRA funds into the system, the size of SEP funding relative to ratepayer expenditures varies considerably by state, but is generally low for the states that have large utility programs, and high for those that do not. Figure 16 displays budget information for rate-payer funded and ARRA programs compiled in a recent Lawrence Berkeley National Laboratory study.²³ The budget for the ARRA efforts includes funding for three major programs: SEP, the Energy Efficiency Community Block Grants (EECBG), and the State Energy-Efficient Appliance Rebate Program (SEEARP). SEP accounts for 47 percent of total funding for these three programs. The right-hand column shows an estimate of annual SEP funding as a percent of annual funding for rate-payer programs. This percentage ranges from 3-10 percent in states such as Massachusetts, New York, and California where rate-payer programs are well-established and well-funded to 35-43 percent in states such as North Carolina, Michigan, and Maine where programs are less well-established and funded.

Figure 16. Ratepayer Program and SEP Budgets for Selected States (\$ millions)

	2009 Ratepayer Program Budgets	PY2009 - 10 Budget SEP, EECBG & SEEARP	Annual SEP Budget/ Ratepayer Budget
California	\$1,367.7	\$314.5	5.4%
Colorado	\$60.0	\$62.9	24.6%
Florida	\$139.8	\$172.3	28.9%
Hawaii	\$35.5	\$38.9	25.7%
Massachusetts	\$208.5	\$25.9	2.9%

²³ Goldman, Charles A. et al. *op. cit.*

Maine	\$20.8	\$38.2	43.1%
Michigan	\$66.2	\$111.2	39.4%
Minnesota	\$73.7	\$62.9	20.0%
North Carolina	\$67.7	\$99.9	34.7%
New York	\$421.2	\$171.7	9.6%
Oregon	\$105.4	\$55.4	12.3%
Wisconsin	\$162.4	\$72.6	10.5%
	\$2,728.9	\$1,226.4	10.6%

Source: Goldman et al. *Interactions between Energy Efficiency Programs funded under the Recovery Act and Ratepayer-funded Energy Efficiency Programs (Draft)*.

- **Programming strategies.** The Lawrence Berkeley National Laboratory study mentioned above closely examines the coordination of ARRA funding applications with rate-payer funded programs in case studies of the development of ARRA applications in 12 states. The authors interviewed 80 “energy efficiency actors” in the case study states, including state energy officials, sponsors of rate-payer funded programs, and local energy efficiency industry experts. The case studies identified the following modes of interaction between the local SEP officials and representative of other programs.
 - o **Inherent coordination** occurs when the state’s public benefits program administrator also administers the ARRA funding, as is the case in New York and Maine.
 - o **Consultation.** This approach occurred when the state energy office consulted with ratepayer-program administrators on current and planned programs in developing their applications for ARRA funding. Several states formally consulted with ratepayer-program administrators, affording an opportunity for exchanging information and learning, but then went their own way and developed programs that targeted similar market segments as ratepayer-funded offerings or occupied very similar programmatic space. For example, in Wisconsin, the state Office of Energy Independence consulted and coordinated appliance rebates with Focus on Energy, the non-profit statewide administrator of ratepayer-funded programs. However, the energy office chose to field its own ARRA-funded program for industrial efficiency. Elsewhere, consultation resulted in closer cooperation in delivery of programs.

- o **Complementary programming.** In these cases, state energy office officials explicitly coordinated with ratepayer-program administrators and designed ARRA-funded programs as complements, enhancements, or extensions of ratepayer programs. In some cases, these programs served different, non-overlapping markets, for example programs that supported residential oil heating savings. In others the SEP PAs supported or enhanced existing rate-payer funded efforts, for example, by providing a web portal to all assistance programs available to customers in a given sector, regardless of program sponsor.
- o **Full collaboration** between the SEO and other program sponsors results from close cooperation in designing and implementing joint programs including comingling and sharing funds, expertise, labor and branding. A few states, California, Hawaii, Maine, Massachusetts, and Minnesota, provide examples of complete collaboration. In Minnesota's Trillion BTU program, the SEO delegated ARRA money to a port authority with more experience in economic development for a revolving loan fund targeting the commercial and industrial sectors. The state's largest utility is adding rebates and engineering assistance for participants. The combined effort is intended to offset nearly all upfront costs for industrial energy efficiency projects.

This range of joint programming approaches, which was obtained to some extent in the pre-ARRA period as well, drives home the importance of taking the specifics of each sampled PA's situation into account in implementing an attribution analysis. For example, under the "inherent coordination" and "full collaboration" scenarios, the individual contributions of the joint sponsors are not visible to the targeted market actors. Therefore, assessment of their relative effects will need to rely of exploration on other data in addition to than market actors' perceptions and response. Similarly, in the case of complementary program, we may need to test the hypothesis that the SEP PA's outreach, publicity and delivery efforts may have encouraged market actors to participate in ratepayer programs, and vice versa, even if those programs were putatively targeted to different populations.

As a first step in addressing the issue of allocation of influence among multiple programs, each PA evaluation will include the development of a map of the program's domain. The map will show, among other things, the identity of other program sponsors in the domain, the activities and offerings of their programs, the levels of resources available, and the duration and scope of the interactions between SEP and other program players and groups of market actors. This map will show whether exploration of the question of multiple program influences is important for the PA in question. It will also help in the formulation of specific research questions to be posed to market actors

in a manner similar to process and logic models commonly used in program evaluations of all kinds.

Influence of ARRA on other program sponsors. One key motivation for assessing the response of other program sponsors to SEP is to provide a basis for allocating attribution among multiple program sponsors other than a simple share of funding. Review of program narratives suggests that, at least in some cases, there might have been no program at all in the relevant domains in the absence of SEP organizational capacity or funding. We also need to account for the possibility that the introduction of large amounts of ARRA funding induced other sponsors to reprogram available resources away from areas served by newly enlarged PAs. Review of narratives of the implementation of SEP prior to and during the ARRA initiative suggest that there were two reasons that sponsors of other programs coordinated with SEP activities. These were as follows:

- **Leverage SEP organizational capacity.** Over its many years of operation, state SEP officials have built up organizational capacity to advance energy efficiency and renewable energy program objectives. This capacity consists of in-house technical expertise, working relationships with regulators, state legislators, state executive officials, business leaders in various sectors and academic institutions, and program delivery capability. All of these could serve the purposes of other program sponsors where mutual advantage could be identified for cooperation with SEP.

To assess whether the availability of SEP organizational capacity affected another sponsor's decisions, we will need to interview representatives of the sponsor to probe the following:

- o Were decision makers in the other sponsors aware of SEP organizational capacity?
- o What were their perceptions of SEP organizational capacity in regard to:
 - Relationships with market actors?
 - Relationships with regulators and other government agencies?
 - Technical expertise?
 - Program delivery capabilities?
- In what specific ways did the other sponsors take SEP capabilities into account in their program planning?

- What do the sponsors of other programs think they would have done if SEP organizational capacity had not been available? For example, would they have changed the objectives of the program, changed the design of the program, changed the volume of program activity and funding, or sought another partner? To what extent are these alternatives consistent with the other sponsors' organizational mission and available resources?

The evaluation team will pose a parallel set of questions to SEP officials who have knowledge of the relevant program history.

- **Leverage SEP Funding.** With the advent of ARRA, SEP was in position to provide significant funding enhancements for programs operated by other sponsors. This money could be used to free up ratepayer resources for use in other market segments, to enhance the level of support available to market actors in the segments already served, or some combination of the two. To assess the effect of SEP funding on the net level of program resources available in a given program domain, we will need to interview representatives of the other programs to probe the following questions:
 - o In what specific ways did other programs take SEP and/or ARRA funding into account in planning their own activities and in allocating total resources to individual programs?
 - o In the absence of the SEP funding, would the other sponsors have allocated the same level of resources to the domain of the SEP PA under evaluation, a lower level of resources, or a greater level of resources?
 - o If a greater level of resources, what elements in the sponsor's portfolio received the funds diverted from the evaluated PA's domain? If a lower level of resources, what elements of the sponsor's portfolio provided the additional funding?
 - o To what extent are the counterfactual funding allocations consistent with the other program sponsor's mission and resources?

5.1.2 Available Methods for Assessing Attribution

Five basic methodological approaches can be found in the energy efficiency program evaluation literature for assessing attribution of savings to specific programs. These include the following.

- **Analysis of self-reports of program effects by targeted market actors (Self-reports).** This approach typically involves surveying samples of actual and/or potential

program participants to elicit their assessment of the program's influence on their decisions to adopt energy efficiency measures or practices. The questions can be structured to probe the effect of the program on the timing, extent, and features of the projects in question, as well as the relative importance of the program versus other decision factors. The responses can then be processed to develop an attribution score using a transparent algorithm.

- **Quasi-experimental designs.** This approach uses well-established quasi-experimental social research designs to assess and quantify program attribution. Common strategies include cross sectional methods that compare the rate of measure adoption in an area or market segment not targeted by the program as a baseline for comparison to rates of adoption in the program area. The difference between the two can be viewed as the program's net effect. Pre-post designs that compare the rate of adoption before and after the program or policy intervention have also been applied, as have mixed pre-post/cross-sectional approaches. Statistical modeling is often used to apply retrospectively quasi-experimental approaches to datasets that describe the response of a group of market actors to a given program. For example, analysis of variance and regression approaches implicitly invoke quasi-experimental designs by estimating program effects while controlling statistically for the effects of other participant attributes such as income, education, facility size, and so forth. Billing analysis to estimate energy savings from program participation is essentially a quasi-experimental approach. In some cases changes in billed consumption over time are compared for participant and non-participant groups. In other cases pooled time series/cross-sectional regression analysis is used to estimate the fixed effects of program participation.
- **Experimental designs.** Experimental design, by which we understand random assignment of eligible market actors to receive different program treatment, provides one of the strongest approaches to assessing attribution. Random assignment directly addresses one of the most serious threats to validity that is inherent in other methods for attributing attribution, namely participant self-selection. Self-selection for participation in voluntary programs generally introduces bias to quasi-experimental analyses because participants often differ systematically from non-participants in factors that affect energy savings that cannot be directly observed and controlled for statistically. Experimental designs have been used recently to evaluate the effect of customer education and information programs. This is a good application of experimental methods because individual participants can be randomly assigned to receive different messages and information products and the marginal cost of program delivery is very low. While evaluation team will look for opportunities to deploy random assignment strategies, we

do not anticipate that they will have much application to this evaluation. Generally speaking, it is necessary to design the delivery of programs to support random assignment. We are aware of few PAs in which this was the case. Moreover, in the absence of the kinds of life and death issues associated with drug trials, it is politically difficult to justify denying access to valuable incentives and services associated with voluntary programs to support evaluation.

- **Price elasticity approaches, including conjoint analysis and revealed preference analysis.** In these two approaches, researchers assess the effect on changes in price on customer's likelihood of purchasing an energy-efficient product or service. The results of these assessments can then be combined with information on the actual effect of the program on the price participants paid for the product or service in question to estimate the effect of a program-related purchase incentive on the pace of sales. In the case of conjoint analysis, customers are asked to rank a structured set of hypothetical products that vary along a number of dimensions, including performance and price. In the revealed preference approach, purchasers are intercepted at the point of sale to gather information on product selection they actually made, its price, and other features.
- **Structured equation modeling.** Structured equation modeling applies a flexible form of path analysis to identify the most likely causal chain from program outputs such as messaging or incentives on the one hand to taking action to adopt an energy-efficient product or practice on the other. Generally, this type of modeling makes use of psychological theories of motivation and action to identify intermediate steps between program stimuli and the desired action. Calibration and testing of these models generally requires survey data from very large samples of market actors. To date, it has been used primarily to assess the effects of information programs.
- **Adoption process models.** One large class of diffusion theories and research rests on contagion models, where the mechanism of adoption is driven by social contact between individuals or firms that have already adopted the technology and those who have not. The most common formulation of the contagion approach is the "mixed influence" model, of which the well-known Bass curve is an example. These models take into account external influences on model adoption, such as prices of alternative products, as well as the pace and density of interactions among those who have adopted the product and those who haven't.

The most well known work in this field, Everett Rogers's *Diffusion of Innovations*. Rogers posits a five-stage sequence that individuals go through the adoption process: knowledge (awareness), persuasion, decision, implementation, confirmation

(evaluation). These stages can be used to structure research on the effects of programs over time. For example, Reed et al. assessed the effects of a program by the Federal Energy Management Program (FEMP) to encourage federal agencies to make use of Energy Service Performance Contracting (ESPC) procedures to implement major energy efficiency improvements in their facilities. To do so, they used periodic surveys of agency employees in position to use ESPC in terms of their adoption stage. Changes in the distribution of the population of targeted employees among the adoption stages were used as indicators of program effects.²⁴

- **Structured expert judging.** Structured expert judgment studies assemble panels of individuals with close working knowledge of the various causes for changes in the market, technology, infrastructure systems, markets, and political environments addressed by a given energy efficiency programs to estimate baseline market share and, in some cases, forecast market share with and without the program in place. Structured expert judgment processes employ a variety of specific techniques to ensure that the participating experts specify and take into account key assumptions about the specific mechanisms by which the programs achieve their effects. The Delphi process is the most widely known of this family of methods.
- **Historical Tracing: Case Study Method.** This method involves the careful reconstruction of events leading to the outcome of interest, for example, the launch of a product, the passage of legislation, or the completion of a large renewable energy project, to develop a ‘weight of evidence’ conclusion regarding the specific influence or role of the program in question on the outcome.

Researchers use information from a wide range of sources to inform historical tracing analyses. These include public and private documents, personal interviews, and surveys conducted either for the study at hand or for other applications.

The historical tracing or case study method provides a great deal of flexibility in dealing with the diversity of program objectives and local conditions that the SEP evaluation will encounter, not to mention the complex issues involved in allocation of attribution and the effect of SEP on other program sponsors’ decisions. However, to be used in assessing attribution, case studies must be rigorously structured and meet standards for documentation laid out in standard guides such as Michael Quinn Patton’s *Qualitative*

²⁴ Reed, John H., Gretchen Jordan, and Edward Vine. *Impact Evaluation Framework for Technology Deployment Programs*. Washington D. C.: U. S. Department of Energy, 2007.

*Research and Evaluation Methods*²⁵ and Miles and Huberman's *Qualitative Data Analysis*.²⁶

Historical tracing relies on logical devices that have been well established historical studies, evaluation of other types of social programs, and legal argument. These include:²⁷

- o Compiling, comparing, and weighing the merits of narratives of the same set of events provided by individuals with different points of view and interests in the outcome.
- o Compiling detailed chronological narratives of the events in question to validate hypotheses regarding patterns of influence. This approach corresponds to quasi-experimental methods that make use of pre/post designs.
- o Positing a number of alternative causal hypotheses and examining their consistency with the narrative fact pattern. This step needs to be taken in every qualitative analysis.
- o Assessing the consistency of the observed fact pattern with linkages predicted by a logic model. This approach is particularly important when cross-sectional and pre/post comparisons are not feasible due to the nature of the program or the content of program records.

In order to control the quality of case studies undertaken for this evaluation, the senior project staff will develop a case study template to be completed by the Lead Evaluator and other appropriate staff. The template will be filled out at the beginning of the research and revised to reflect the inevitable mid-stream adjustments that will need to be made as evidence is sought and sifted. The template will include the following:

- o Clear statement of hypotheses to be tested.
- o Statement of methods and approaches to be used in testing those hypotheses, including statement of researchable issues.

²⁵ Quinn-Patton, Michael. 2002. *Qualitative Research & Evaluation Methods, 3rd Edition*. Thousand Oaks, CA: Sage Publications.

²⁶ Miles, Matthew B and A. Michael Huberman. 1994. *Qualitative Data Analysis, Second Edition*. Thousand Oaks, CA: Sage Publications.

²⁷ See Miles & Huberman op. cit. pp 245 – 280 for an exhaustive list of analytical tactics that can be applied to identify and test conclusions that can be drawn from quantitative data.

- o Listing of SEP-informed organizations to be included in the case and their roles in the program.
- o Listing of SEP-informed individuals to be interviewed and the objectives of those interviews.
- o Listing of documents to be reviewed and their expected contribution to the case study.
- o Identification of the basic methods to be used in assessing causal relationships between program activities and outcomes.

KEMA will review the template with ORNL and its advisors and offer suggestions for its revision prior to the inception of research.

In order to contribute valid and useful information to an attribution assessment, any respondent to in-depth interviews needs to have extensive knowledge of the market and policy domain under study, as well as some awareness of the SEP programmatic activities under evaluation. Prior to conducting in-depth interviews as part of the research to support any approach to attribution analysis, the evaluation team will administer a screener to the potential interviewees to ensure that they meet those criteria. Those who do not meet those minimum criteria will not be interviewed to assess attribution. They may be interviewed for other purposes, such as to develop information on baseline practices or other key aspects of market characterization.

5.1.3 Application of Available Methods to Evaluation PAs in Different Groups

Figure 17 displays the Evaluation Team's assessment of the applicability of the attribution assessment methods discussed above to the key research questions by PA grouping as established in Section 4, where those distinctions are meaningful. In the sections we propose an approach to attribution assessment for each of the program groups that appear in Figure 17. We also provide our rationale for the selection of methods in those sections.

Figure 17. Applications of Attribution Assessment Methods to Evaluation of PAs in PA Groupings

- ☐ = Primary Attribution Analysis Approach
- ☐☐ = Secondary Attribution Analysis Approach
- = Tertiary Attribution Analysis Approach

RESEARCH QUESTION/PA Groups	Participant self-reports	Quasi-Experiments	Experimental Designs	Price Elasticity	SEM & Path Analysis	Adoption Process	Expert Judging	Case Studies
#1. MARKET ACTOR RESPONSE								
Building Retrofit and Equipment Replacement	☐	☐	---	☐	---	☐	---	☐
Renewable Energy Market Development – Projects	☐	☐	---	☐	---	☐	---	☐
Renewable Energy Market Development – Manufacturing	☐	---	---	---	---	---	☐	☐
Clean Energy Policy Support	☐	---	---	---	---	---	☐	☐
Information and Training Programs	☐	☐	☐	---	☐	---	☐	☐
Codes & Standards	---	---	---	---	---	---	☐	☐
#2. INFLUENCE OF OTHER PROGRAMS								
Building Retrofit and Equipment Replacement	☐	☐	---	☐	---	☐	---	☐
Renewable Energy Market Development – Projects	☐	☐	---	☐	---	☐	---	☐
Renewable Energy Market Development – Manufacturing	☐	---	---	---	---	---	☐	☐

**Figure 17 (continued): Applications of Attribution Assessment Methods
to Evaluation of PAs in PA Groupings**

RESEARCH QUESTION/PA Groups	Participant self-reports	Quasi-Experiments	Experimental Designs	Price Elasticity	SEM & Path Analysis	Adoption Process	Expert Judging	Case Studies
Clean Energy Policy Support	---	---	---	---	---	---	□	□
Information and Training Programs	□	□	---	---	□	---	---	---
Codes & Standards	---	---	---	---	---	---	□	□
#3. SEP INFLUENCE ON OTHER PROGRAMS								
All PA Groupings	---	---	---	---	---	---	□	□

- = Primary Attribution Analysis Approach
- = Secondary Attribution Analysis Approach
- = Tertiary Attribution Analysis Approach

5.2 Attribution Approach 1: Building Retrofit and Equipment Replacement, Renewable Energy Market Development – Products, Information and Training Programs

In this section, we present details on the proposed methods to address the fundamental research questions in the order in which they are discussed in Section 5.1.

5.2.1 Assessment of Market Actor Response

As discussed above, “first order” assessment of what targeted market actors’ response to the program is assessed by characterizing what they would most likely have done in the absence of SEP services. For PAs in the Building Retrofit and Equipment Replacement group, this assessment will rely most heavily on data gathered directly from participants. Evaluators of ratepayer funded programs in the New England states, California, Wisconsin, and other jurisdictions have developed standard question sequences to characterize the extent to which individual sample participants in various types of programs were “free riders” and the extent to which they were influenced to by the program to undertake related energy efficiency improvements without further assistance (spillover). The free ridership and spillover scores for individual participants are then aggregated to develop a “net-to-gross” ratio which is applied to estimates of savings impacts to estimate net savings for the program.

Most of these questions explore a number of dimensions of program effects, including those on timing of retrofit and replacement projects (acceleration), quantity of measures installed, and efficiency level for those technologies in which a number of efficient variants exist. These sequences and scales offer some advantages in that they have been used frequently for IOU incentive programs and their performance in the IOU utility program evaluation field is understood. On the other hand, many of them contain numerous questions and consistency checks. The KEMA evaluation team will work with ORNL and its advisors to adapt these sequences to the current project, with the objective of developing a few short, widely applicable sequences that impose minimal respondent burden.

Based on our review of SEP activities in the PY2008 and ARRA periods, we anticipate that we will encounter the following special cases and potential exceptions to this approach.

- **Multiple decision makers or observers for large projects.** Many PAs in the Building Retrofit BPAC involve the provision of significant funding to a small number of large

projects. In these cases, mischaracterization of the level of free ridership for a single project can seriously affect the total attributed savings for the PA. In such cases, it will be best to interview more than one decision maker or observer for each sampled project and ensure the respondent is an appropriate source. For example, the Wisconsin approach calls for interviews with informed suppliers as well as facility owners in cases where owner report low program influence and high supplier influence. Similarly, public sector facilities managers may have a very different view of what a state or municipal government would have done in the absence of the program compared to the relevant capital budget official. It may be appropriate to obtain information from both or to weight the opinions of the financial decision maker more heavily, or to discard the less informed opinion. For particularly large projects that account for 25 percent or more of the savings for a given PA, it may be worthwhile to explore program influence using a case study approach with multiple informants. Of course, this approach would only apply in high rigor studies.

- **Use of cross-sectional analysis for selected programs.** Generally speaking we do not believe that it will analytically useful or cost-effective to develop data for cross-sectional analyses that compare the pace of measure adoption between participant and non-participant groups or between areas served by a particular program type and those that are not. Most of the programs under review are offered statewide, and the range of “non-program” areas available is shrinking for many types of measures. On the other hand, if we find that a large number of states are offering programs that support a given technology or service approach, this may offer the opportunity to structure a comparison group in the states that are not. This could then support a number of different PA evaluations.
- **Inclusion of vendor surveys.** Many of the PAs in the Loans, Grants, and Incentives BPAC are designed such that vendors play important roles in program delivery and marketing. This is particularly the case for “deep retrofit” programs in the residential sector, many of which are based on the Home Performance with Energy Star model. This approach requires that vendors invest in training and equipment for home energy diagnostics and commit to delivering home improvements guided by those diagnostics. It will be necessary to conduct interviews with vendors to assess the effects of the program on their business and professional practices in order to characterize the attribution of savings to the program.

Figure 18 displays potential sources of information on to characterize program effects to market actors for the PAs in the BPACs included in this group. Deployment of these approaches in the

evaluation of a given PA will depend on the rigor level and the specifics of the PA's implementation.

Figure 18. Overview of Research on Market Actor Effects

	Non-Residential						Residential	
	Building Retrofit			Loans, Grants, Inc			LGI	
	Owner	2 nd Owner Ren	Vendor	Owner	2 nd Owner Ren	Vendor	Owner	Vendor
KEY RESEARCH QUESTIONS								
Basic Attribution Questions								
• Timing of installations in the absence of the program	☐	☐	☐	☐	☐	☐	☐	
• Quantity or extent of installations in the absence of the program	☐	☐	☐	☐	☐	☐	☐	
• Efficiency level of equipment in the absence of the program	☐	☐	☐	☐	☐	☐	☐	☐
• Other potential influences on the decision making process	☐	☐	☐	☐	☐	☐	☐	
• Relative importance of the program versus other influences	☐	☐	☐	☐	☐	☐	☐	☐
CONTEXT AND CONSISTENCY								
• Past levels of adoption of the technologies in question	☐	☐	☐	☐	☐	☐	☐	☐
• Barriers to adoption of the technologies in question	☐	☐	☐	☐	☐	☐	☐	☐
• Internal resources available to identify opportunities and manage projects	☐	☐		☐	☐		☐	
• Prior understanding of benefits and costs of adoption – information effects of the program	☐	☐	☐	☐	☐	☐	☐	☐

☐ = Important Source

☐☐ = Potential Supplementary Source

5.2.2 Relative effect of multiple programs

Assessment of the relative effect of multiple programs and assignment of attribution credit among them will proceed in the following steps.

1. **Determine whether there were significant programs targeting roughly the same market actor responses or goals that were operating in the same domain as the**

PA under evaluation during the program period. Such programs could be offering the same kinds of assistance to the PA's targeted market sector, for example: a ratepayer program that provided incentives for retrofit projects in commercial buildings. They could also include programs that promoted similar measures to other groups in the domain; for example, training in proper HVAC specification and installation for commercial contractors. Such information would be gathered through interviews with state energy officers, review of web sites for local ratepayer programs, and contacts with local program sponsors to verify offerings.

2. **Characterize the operating relationship between the programs.** This step will be important for shaping the questions to market actors regarding their perceptions and use of the other programs as well as for assessing the consistency of different hypothetical causal chains with the facts of the case. The potential forms of the relationship are as follows.
 - The operators of the various programs essentially make no mention of the others, leaving it entirely up to market actors to integrate the various services in completing projects.
 - The operators of the various programs notify participants of the availability of the other programs and of their potential applications, but did not take active steps to promote integration, such as providing the market actors with application materials.
 - The operators of the various programs actively cooperate in promoting participation in both programs, for example by providing and tracking referrals or expediting applications from participants in the cooperating programs.
 - The programs are jointly administered under one brand.
 - Potential cost-sharing between organizations/programs.
3. **Characterize the relationship between the offerings of the programs from the point of view of market actors.** It will be important to understand whether the programs are competing for some part of the participants' value chain in implementing energy efficiency or renewable energy measures or are offering complementary services. For example, the analytic treatment of customer responses to two programs that both offer financial incentives will be different from the treatment of two programs, one of which offers energy audits and technical assistance to owners while the other offers financial incentives.

4. **Identify and characterize other potential influences that affect the targeted responses of market actors in the PA's domain.** These may include legislation and regulations affecting the development of various kinds of projects, for example environmental and interconnection rules related to combined heat and power installations, changes among firms in the local supply chain of the relevant technologies, past and current market preparation and change acceptance efforts, and so on.
5. **Develop a map of the program's domain.** The PA evaluation team will combine the results of the first four steps into a map of the program's domain. It will show the identity of major public and private sector organizations active in the domain, their missions or organizational interests, and the range of their activities. For all programs, including the PA under evaluation, the map will show the services offered, eligibility requirements for participants and projects, the total funding available in the current year, and the likely duration of funding and activity. This map will serve a number of key functions in the analysis, including:
 - Provide a reference point for tailoring of the attribution question sequences for market actors.
 - Provide a reference for development of the process model of the program, if such is needed for the attribution analysis.
 - Provide inputs to the analysis of causal links between program activities and outcomes.
 - Support allocations of attribution credit among programs recommended by the Lead Evaluator.
6. **Collect information on relative program influence from SEP informed market actors.** Key data to be collected from the targeted market actors will include the following:

Facility and Home Owners

- What percentage of facility or home owners who implemented projects with the support of other programs were aware of those programs prior to participation in the SEP PA?
- Did those owners become aware of the other programs through SEP PA activities, through other channels, or through both? Did they become aware of the SEP PA through other programs? In all cases we will design the

relevant questions to assure that the programs under evaluation are named in such a way that the respondent is most likely to recognize them.

- If the owners had not taken part in the SEP PA, how likely is it that they would have participated in the other program?
- Did the owner become aware of the opportunities and costs associated with the measures in question through the SEP PA, through other programs, through other channels, or through some combination? Through which channel did they first become aware of the opportunities?
- If owners were aware of the opportunities offered by the measures prior to participation, what specific barriers prevented from undertaking them? What is the relationship between these barriers and the offerings of the various programs?
- What relative level of importance or influence do SEP PA participants assign to the various programs in their decision to adopt the technologies in question?

Supply-side Market Actors. As discussed above, it will in some cases be useful to gather information from SEP-informed designers, engineers, and installation contractors involved with larger projects to assess attribution. Some key items of information from these sources include the following:

- What specific barriers were preventing the owner from undertaking the project in question? Which of these had been most decisive?
- What were the contributions of the various programs to the completion of the project?
- What was the relative importance of those contributions to the completion of the project?

7. **Develop and apply a scoring algorithm to the market actor data.** Given the diversity of PA activity in the Building Retrofit and Equipment Replacement group portfolio and the variety of relationships between PA and other local program sponsors, we do not feel it is possible or appropriate to develop a single, universally applicable scoring algorithm for responses to the market actor questions listed above. Figure 19 displays the market actor data to be taken into consideration in developing individual scores and our assessment of the relative importance of those data in developing such scores. Each score for individual sample projects supported by the PA under evaluation will represent

the percentage of savings for that site to be attributed to the PA. The scores will be weighted by the site level savings where appropriate and aggregated to the PA level. All scoring systems and methods used to estimate net-to-gross ratios or to otherwise characterize program influence will be reviewed with ORNL and its advisors prior to their application in the study.

Figure 19. Considerations in Scoring Market Actor Data on Relative Importance of Multiple Programs

SCORING FACTORS FOR RELATIVE ATTRIBUTION OF PROGRAM INFLUENCE	Non-Res		Res
	Building Retrofit	Loans, Grants,	Loans, Grants,
FACILITY OR HOME OWNERS			
<ul style="list-style-type: none"> Channel through which owners first became aware of the relevant measure 	☐	☐	☐
<ul style="list-style-type: none"> Channels through which owners became aware of the various programs; sequence of referrals or applications. 	☐	☐	☐
<ul style="list-style-type: none"> Nature of barriers to implementation of relevant measures in relation to program offerings 	☐	☐	☐
<ul style="list-style-type: none"> Rating of importance of different programs in the decision to proceed with implementation of the measure. 	☐	☐	☐
<ul style="list-style-type: none"> Importance of factors other than programs in the implementation decision 	☐	☐	☐
VENDORS			
<ul style="list-style-type: none"> Observed barriers to implementation in relation to the offerings of the various programs. 	☐	☐	☐
<ul style="list-style-type: none"> Relative importance of the observed barriers. 	☐	☐	☐
<ul style="list-style-type: none"> Assessment of the relative contributions of different programs to the owners' implementation decisions 	☐	☐	☐
<ul style="list-style-type: none"> Assessment of the importance of factors other than the programs in the implementation decisions 	☐	☐	☐

☐ = Important Consideration in Scoring

☐☐ = Supplementary Consideration

8. **Adjustment of aggregate attribution scores for qualitative considerations.** Given the nature of the SEP PAs and their operating environment as described above, we

believe it will be necessary to adjust the aggregated market actor score for information that is available to the evaluators but not to market actors and that affects the logic of the evaluation. Examples of such adjustments will include the following.

- When program offerings are complementary, such as audits from one and financial incentives from the other, it will be necessary to accord at least a minimum level of attribution to both in accordance with the logic of the situation, even if the results of the market actor survey accord little or no weight to one or the other. On the other hand, in instances where it is up to the market actors to integrate the services of various programs on their own, it may not be appropriate to adjust aggregated scoring algorithm results.
- In situations where programs are offered jointly, it may make more sense under the logic of the situation to allocate attribution credit according to funding than by the aggregated results of the market actor scoring.

5.2.3 Influence of SEP on Other Programs

See Section 5.1.1 for a discussion of questions to be addressed in assessing the influence of SEP on the resource allocation decisions of other program sponsors in the sampled PAs' domains. Information to address these questions will be gathered from the following sources:

- In-depth interviews with state energy office officials.
- Interviews with managers of the other major programs in the sampled PA's domain.
- Review of relevant documents, including regulatory filings, program plans.

These research activities will be undertaken early in the PA evaluation, prior to any surveys or interviews with market actors. In that way the results of our assessment of the influence of SEP activities on other programs can be integrated into the analysis of data received from the market actors. Based on review of information from the sources listed above, the Lead Evaluator and staff assigned to the sample PA will make a determination of the extent to which SEP influenced the allocation of program resources by other sponsors. The Lead Evaluator will summarize this determination and the evidence for it in a memorandum which will become part of the report for that PA evaluation.

5.3 Attribution Approach 2: Renewable Energy Generation and Capacity – Manufacturing, Clean Energy Policy Support, Codes and Standards

In this section we describe the methods to be used in assessing attribution for PAs in the Renewable Energy Market Development – Manufacturing, Clean Energy Policy Support, and Codes and Standards BPACs. We anticipate using a combination of these methods for assessing attribution for PAs in the Clean Energy Policy Support BPAC. However, given the highly variable nature of those programs it is not possible at this time to specify the attribution approach in detail.

In the next two sections we first describe the mechanisms by which we anticipate that the PAs in this group will accelerate adoption of energy efficiency and renewable energy technologies. We then describe the range of methods we plan to use to quantify the effects of those mechanisms and their application to evaluation of PAs in the BPACs included in this group.

5.3.1 Renewable Energy Market Development – Manufacturing

Mechanisms for accelerating adoption: changes to product attributes and availability.

The theory linking financial support for product enhancement and expansion of manufacturing capacity posits that such support will increase the attractiveness of renewable energy systems to purchasers by reducing their price, increasing their performance, and adding features that reduce the costs or performance uncertainty of the targeted technology or the market in which that technology operates. This phenomenon has been extensively studied under the rubric of learning effects, and a number of studies have identified significant learning effects from government support for wind and solar technologies.²⁸

Thus, if the manufacturing-oriented programs in the SEP portfolio work as expected, we should expect to see the one or more of the following over the years following the disbursement of funds:

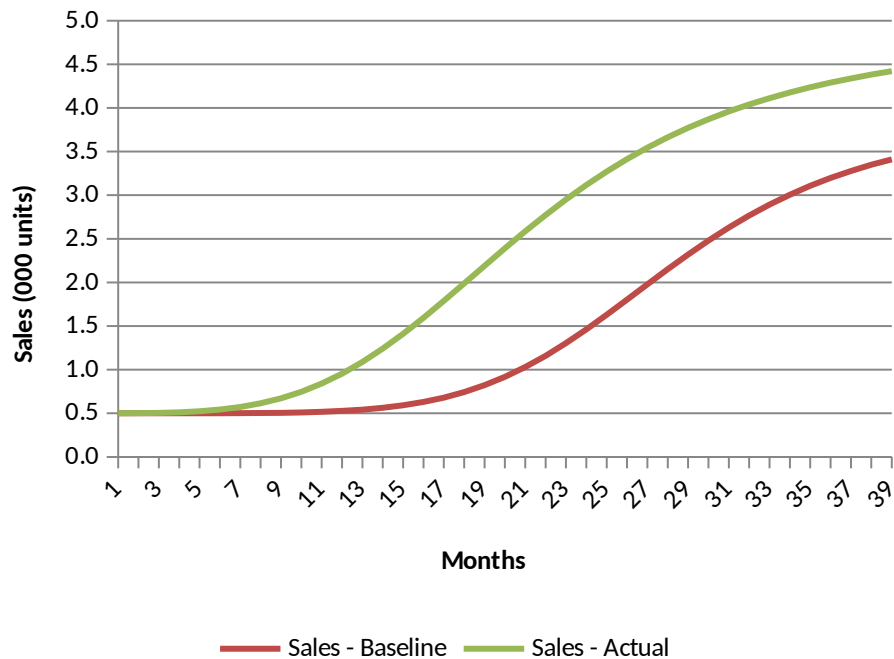
- Reduction in the price per unit of capacity for the equipment in question.
- Moderation in the price per unit compared to non-SEP influenced markets.

²⁸ See, for example, Jako, P. *Learning and Diffusion for Wind and Solar Power Technologies*. Petten, NL: Energy Centre of the Netherlands. 2002

-
- Improvement in market or technical (non-energy) performance of the technology, such as a reduction of down-time, lower maintenance cost, easier accessibility, reduction of time for parts delivery, etc.
 - Increased level of market acceptance or the reduction of a market barrier.
 - Increased renewable energy generation per unit of capacity for the equipment in question.
 - Stabilization in the renewable energy generation per unit of capacity for the equipment in question compared to non-SEP influenced markets.
 - Increase in sales of the equipment in question beyond a baseline that represents the sales forecast for the corresponding equipment prior to the influx of SEP resources into the product development and manufacturing process.
 - Stabilization of sales during a period of market decline caused by non-SEP changes in the market (increases in manufacturing costs, decreased federal incentives, changes to state rules, policies and regulations, changes to renewable energy standards, changes to utility cost allocations to renewables, etc. etc.)

As part of the process of developing the study plans for individual PAs in this BPAC, we will develop indicators of market acceleration that are relevant to the technologies and markets involved. Given the timing and duration of the SEP activities and of this evaluation, it is likely that the full effects of the support to manufactures will not work its way through the sales cycle by the end of the evaluation period. Therefore, it will be necessary to forecast actual sales with the SEP assistance in place, as well as baseline level sales that posit no assistance from SEP. These forecasts would be based on in-depth interviews with industry experts who are aware of SEP and the various factors that are driving change in their markets. Figure 20 shows an example of this kind of dual forecast. Energy impacts attributable to the SEP PA would correspond to the average renewable energy generated per unit of capacity multiplied by the capacity of the net units sold per this analysis.

**Figure 20. Forecasts of Measure Sales:
Baseline and Actual with Program Support**



Research and analysis methods. Given that most of the PAs in this group were funded by ARRA, we do not anticipate being able to afford the full apparatus of an Expert Judging process. We will, however, prepare a fact package for the experts consulted in order to provide support for their sales forecasts. These fact packages will include information on the following, drawn from interviews with project principals and secondary sources.

- Total market size (unit sales per year) for the products under evaluation.
- Current costs per unit of capacity: average and range.
- Past trends in costs and performance per installed capacity, forecasts of those trends, drivers of the forecasts.
- Current and anticipated near term changes in other market drivers, including interconnection, feed-in tariffs, carbon markets, conventional energy prices, etc.
- Current sales of the product in question and related products made by the same manufacturer.

After providing the industry experts with this fact package, we will ask them to provide forecasts of annual unit sales under baseline (no program) and actual conditions. We will probe the following in regard to the sales forecasts.

- Perceptions of the competitive position of the supported products prior to the program.
- Perception of changes in the competitive position of the product as a result of SEP support.
- Perception of changes in competitors' products over the same period.
- Expectations concerning the development of other market drivers, including interconnection, feed-in tariffs, carbon markets, conventional energy prices, etc.
- Potential improvements to product performance or cost effectiveness that may be opened up by the innovations funded by SEP.
- Perceptions of factors affecting the ability of the subrecipient to respond to continuing changes in the market and the competitive environment, such as capitalization and marketing capabilities in relationship to peer companies.

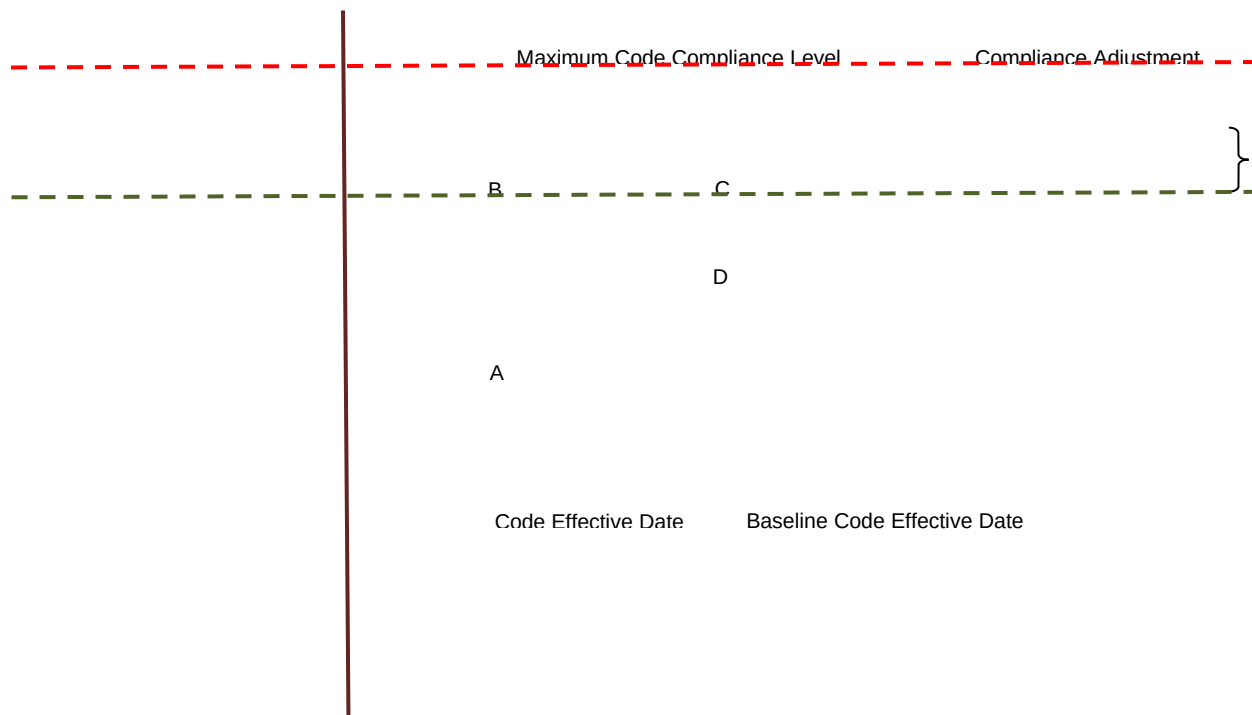
5.3.2 Codes and Standards Programs

Mechanisms for Accelerating Measure Adoption. The PAs in the Codes and Standards group feature two principal mechanisms for accelerating the adoption of energy-efficient designs and equipment in new construction and major renovation projects. These are:

- **Training for enhanced code compliance.** These programs consist of efforts to train state and local code enforcement officials in enhanced code compliance methods, which include improved field inspections and organization of building inspection management processes at the municipal and state level. These enforcement enhancement initiatives account for most of the 43 PAs in the Codes and Standards BPACs for PY2008 and the ARRA period.
- **Acceleration of code adoption.** As discussed in Section 4, an important element of SEP for accelerating code adoption is the requirement in the ARRA Funding Opportunity Announcement that required the governor of each state receiving SEP-ARRA funding to undertake to adopt the most advanced residential and commercial energy codes currently in use and to take steps to assure compliance.

Figure 21 shows the savings to be obtained by acceleration of code adoption, which is represented by the area bounded by points A, B, C, and D. Savings are increased to the extent that compliance deficits can be reduced, which would raise the dashed line towards the maximum code compliance level.

Figure 21. Acceleration of Code Adoption



Research and Analysis Methods. The main objectives of research and analysis for attribution analysis of codes and standards programs will be to:

- Develop a forecast of the baseline year in which the state in question would have adopted the IECC 2009 residential building code and a commercial code that meets the ANSI/ASHRAE/IESNA Standard 90.1–2007.
- Develop a forecast as to when compliance efforts achieve a 90 percent compliance rate.

For the high rigor evaluations of Codes and Standards PAs sampled from PY2008, the KEMA team will use structured SEP informed expert judging methods, such as the Delphi process, to gather, organize, and analyze information and opinions in regard to the two analytical objectives listed above. For the medium-high-rigor studies Codes and Standards PAs sampled from the

ARRA period, we will use structured in-depth interviews with experts on building code adoption and enforcement in the states from which the programs were sampled.

The steps in this research will be similar for both the high- and medium-high-rigor approaches, except as noted. They are:

- **Identification and screening of experts.** Generally, we will attempt to assemble a panel of experts from each sampled state who bring to their assignment a diversity of views and experience on the questions at hand. Specifically we will attempt to identify and interview or recruit onto the expert panel representatives of the following kinds of organizations: state energy office managers with responsibility for code efforts; state agencies with responsibility for code adoption and support of local building departments in code enforcement; local building department officials; building inspectors; general contractors in the commercial construction industry; home builders; residential and commercial property developers; construction market observers from industry publications; consulting firms; or academia. The judges will be screened to ensure that they do not have a financial or professional interest in the outcome of the assessment of the program's effects. We will also pose questions concerning the prospective judges' recent experience in the markets under evaluation to ensure that they are well-positioned to have observed changes in the market and to have access to specific information concerning the effects of code changes. They should also have direct observations of the role of SEP program activity on the changes made to the code.
- **Reducing disparity among experts in knowledge of the product, the market, and the project.** Experts enter the assessment process with distinctly different levels of knowledge and understanding of the SEP program's influence on the product, the market, and the operation of the project. These differences can make it difficult to bring the full range of their knowledge and experience to bear on their forecasts. To address this issue, researchers generally prepare fact packets that detail conditions in the market to be assessed, relevant regulations, and so forth. The KEMA team will identify informed market actors and develop these fact packets. We anticipate that they will include information on the following:
 - o Volume and mix of construction by building type and size over the past several years.
 - o Administrative mechanisms by which code changes are introduced and adopted.
 - o Identification of the SEP role in that effort
 - o Narrative of the technical and political aspects of the most recent code changes.

- o Administrative structure of code enforcement.
- o Current levels of code enforcement activities, including levels of staffing and funding at the state and municipal levels.
- o Schedules for subsequent code adoption proceedings prior to the agreements to change the codes entered into as part of the ARRA process.

Much of this information will need to be collected in any case to support estimates of energy impacts discussed in Section 4.

- **Clarification of assumptions concerning drivers of market acceptance.** A judge's view of the likely trajectory of market acceptance will depend on his or her assumptions about trends in drivers of market acceptance, such as the price of energy or changes in construction costs, information flows on products and services, market acceptance barriers and how these are overcome, etc. Researchers have used a number of approaches to clarify these assumptions. One is to provide judges with a number of scenarios concerning the development of drivers over the forecast period, and to request that the judges provide forecasts under each scenario. A second is to probe the judge's rationale for the forecast in a follow-up round of questioning. These scenarios will be included as part of the fact packet.
- **Initial round of questioning.** Based on the information described above, the SEP informed experts will be asked to forecast the year in which the state in question would have adopted the relevant codes in the absence of the SEP initiatives that have occurred in their state, and to provide the rationale for that forecast. Similarly, judges with experience in code enforcement activities in the state will be asked to predict the highest level of code enforcement the state would have been able to reach on its own (without ARRA or SEP) in terms of percentage of projects or square footage of new construction completed and to provide their rationales for those answers.
- **Iterative rounds to increase reliability (high rigor only).** The first round of forecasts usually yields a broad range of predictions – too broad to be viewed as a reliable guide to the future. To increase the reliability of the forecast, researchers typically conduct at least a second round of inquiry and, in some cases, additional rounds. In these rounds, the individual SEP-informed judges are shown the average values of the forecasted indicators. They are asked to provide the rationale for their forecasts and are offered the opportunity to revise the forecasts. This process generally yields a tighter distribution of the forecasted variables, although outliers are seldom eliminated entirely.
- **Aggregation of results.** There are a number of methods available for aggregating the results of these kinds of data collection activities. One is to take average values of key

parameters, such as the elapsed time from the present day for the baseline code effective date. Statistical measures of reliability are available for results based on averages which can be used to characterize the consistency of experts' judgments. On the other hand, averaging methods may give undue weight to extreme values, especially where only a few individuals are involved. It may therefore be better to use non-parametric estimators, such as the median to represent the aggregated results of the expert interviews or judging.

- **Assessment of corroborating evidence.** For both high and medium high rigor studies, we will interview SEP-informed officials, including SEP Managers, directly responsible for overseeing code adoption and enforcement activities at the state level to ask them the same questions posed to the panel of experts. These judgments will be weighed against the results of the methods described above in developing a final set of net-to-gross ratios to be applied.

6. Evaluation of Carbon Impacts

6.1 Assessment of Carbon Impacts

The assessment of gross carbon dioxide (CO₂) savings will be done for each broad program category and for the individual indicator activities. Annualized CO₂ reductions achieved as a result of SEP-funded efforts will be calculated and reported for each year over the effective useful lifetime (EUL) of the measures evaluated. When the consumption of energy from fossil fuel resources is reduced, the CO₂ emissions that would have resulted from burning those fuels are avoided. Likewise, when renewable energy is used as an alternative to fossil fuels, the CO₂ emissions associated with the replaced fuels are avoided.

In this study, the carbon emissions avoided from SEP-funded energy efficiency and renewable energy activities will be reported nationally and for each state. The assessment of gross CO₂ savings will be done for each broad program category and for the individual indicator activities. Annualized CO₂ reductions achieved as a result of SEP-funded efforts will be calculated and reported for each year over the effective useful lifetime (EUL) of the measures evaluated.

The approach to be taken is consistent with recommendations contained in the Model Energy Efficiency Program Impact Evaluation Guide²⁹ (“the Guide”). As noted in the Guide: “The methods for determining avoided emissions values for displaced generation range from fairly straightforward to highly complex. They include both spreadsheet-based calculations and dynamic modeling approaches with varying degrees of transparency, rigor, and cost. Evaluators can decide which method best meets their needs, given evaluation objectives and available resources and data quality requirements.”

For this study, the basic approach selected employs the use of emission factors as follows:

$$\text{avoided emissions}_t = (\text{net energy savings})_t \times (\text{emission factor})_t$$

The emission factor is expressed as mass per unit of energy (e.g., pounds of CO₂ per MWh), and represents the characteristics of the emission sources displaced by reduced generation from conventional sources of electricity.

²⁹ National Action Plan for Energy Efficiency (2007). *Model Energy Efficiency Program Impact Evaluation Guide*. Prepared by Steven R. Schiller, Schiller Consulting, Inc. <www.epa.gov/eeactionplan>

A number of options exist for selection of emission factors. Non-baseload emissions rates from the U.S. Environmental Protection Agency's (EPA) Emissions & Generation Resource Integrated Database (eGRID)³⁰ can be used to quantify avoided emissions. Non-baseload emission rates have been developed to estimate the emissions from marginal generation units, which are those most likely to be displaced by energy efficiency and/or renewable energy programs and projects. Although the non-baseload emission metric is recommended by EPA for this purpose,³¹ the contractor team will use a higher rigor process for estimating carbon emissions. This will require the following three step process if sufficient libraries of load shapes and corresponding emission rates are available:

- Develop a appropriate regional (or State if available) library of load shape profiles by BPAC;
- Distribute savings across three recommended periods: summer peak, summer off-peak, and everything else; and,
- Apply emission rates roughly corresponding with the region represented by the most appropriate load shape profiles in the library to capture baseload and non-baseload generation dispatch schedules.

To ensure reasonable national coverage for a portfolio of BPAC measures, building or end use load shape data will need to be developed from various sources (e.g., KEMA, Itron, other sources) and blended and weighted from a variety of end uses or building types to be representative of the BPAC. Additionally, load shape data and emission rates will need to be developed in tandem to ensure that regional/state representation is roughly similar between the load shape profiles and the region it represents.

³⁰ eGRID2010, the most recent version will be used for this analysis.

³¹ E.H. Pechan & Associates, Inc., "The Emissions & Generation Resource Integrated Database for 2010 (eGRID2010) Technical Support Document," Prepared for the U.S. Environmental Protection Agency, Office of Atmospheric Programs, Clean Air Markets Division, Washington, DC, December 2010.

6.2 Presentation of Results

KEMA will develop national and state-level net energy savings and renewable generation estimates using the methods described in Sections 3, 4, and 5. The state-level estimates will be based on the portfolio of evaluated programmatic activities supported by that state. We will then estimate carbon savings associated with all energy savings and/or renewable generation for all relevant programmatic activities within each state to estimate state-level avoided emissions and aggregate up to a national level estimate.

7. Evaluation of Employment Impacts

7.1 Assessment of Employment Impacts

7.1.1 Broad Parameters of Jobs Assessment

The measurement of annual job impacts will occur at the state-level for each BPAC. Those BPACs containing several heterogeneous program activities will require job impact estimation for each of those. This assessment will be developed for the pre-ARRA program year (2008) and each ARRA program year (2009, 2010, and 2011). The result of each assessment is a time-series of annual job impacts resulting from the short-term incremental spending related to projects (within the PA/BPAC/the SEP program) and the longer-term effects from net (verified) energy savings, and any market transformation that results. The short-term and long-term effects of program year activities on which job impacts will be gauged are associated with the respective sample definitions as described in Section 3 above, and the associated data collection efforts, described in Section 4 above.

7.1.2 Economic Impact Model for identifying Job Impacts

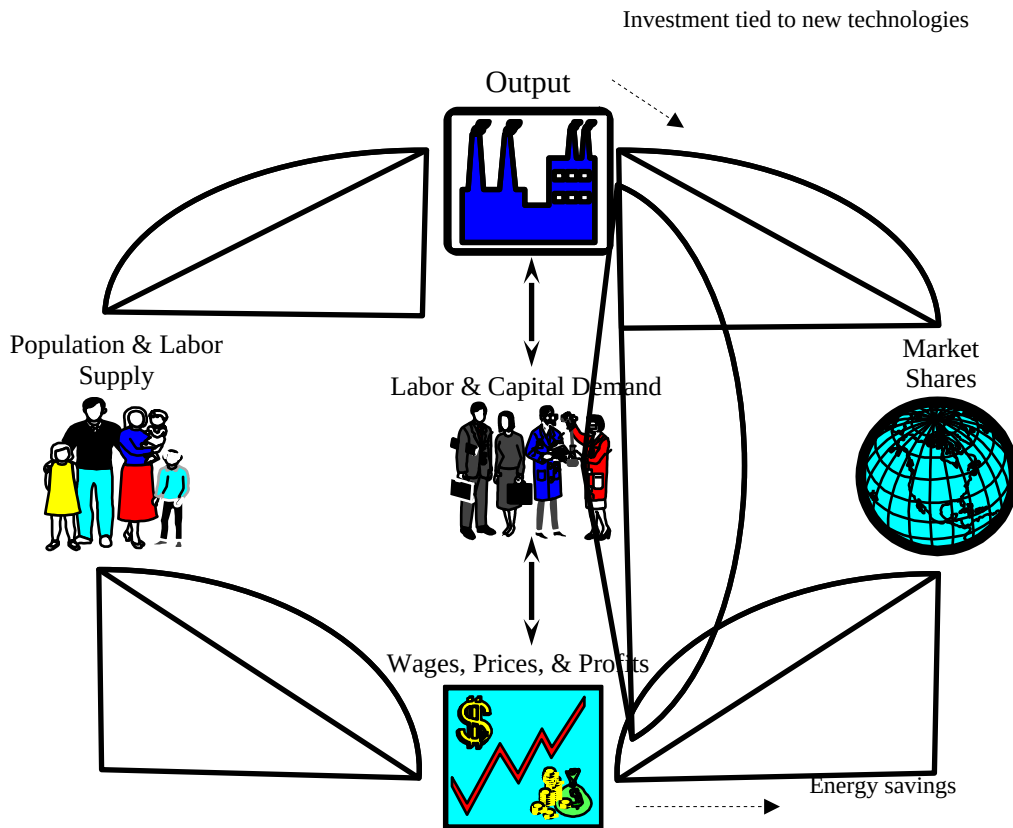
Our proposed approach includes a 51-region (state) REMI *Policy Insight* simulation model. Information describing the short-term and long-term project-related effects will be introduced into this economic model to identify the annual projection of job impacts. This analysis system has been applied to numerous energy and environmental policy/program analyses, some applications specifically within evaluation activities³². A brief overview of the REMI model capabilities follows below.

This model is chosen over others since it has the relevant economic levers and feedbacks to handle the types of effects expected to flow from such project spending and energy saving (generating) technology adoption. The model is a computable, general equilibrium (CGE) simulation forecasting system of industry-level activity for 70 different industries (approximating three-digit NAICS definitions of business activity) through the year 2050. It is well-specified through its internal logic or equation set, such that feedbacks among economic stakeholders (households and businesses) are captured when more energy-efficiency and renewable generation investments take place. Figure 22 portrays the basic concept of what the REMI

³² Wisconsin *Focus on Energy* Biennial Economic benefit Evaluations (2002 through 2010), Connecticut Long-term Sustainable Solar Strategies, Macro-economic Impacts from All Cost-effective energy-efficiency for New England; Evaluation of CCEF OSDG and Small Solar Programs.

model captures for a region's economic impacts (a region can be a county/state or any combination of county building blocks). There are five major blocks to a region's economy (e.g. Output, Labor & Capital Supply, etc.), each block contains numerous equations, and the arrows depict the feedback between different components of an economy. In a multi-state model (of 51 regions) one can envision 51 economies such as in Figure 22 which will also exhibit feedback between other states (inter-regional) for labor flows (commuters) and trade in manufactured goods and in services. Unique to the REMI model among the class of competing regional economic impact frameworks available is the linkage to the market shares block. Policies or investments that change the underlying cost-of-doing business for an industry in region k will affect that industry's relative competitiveness (relative to the U.S. average for that industry) and its ability to retain/gain sales within its own region, elsewhere in the multi-region marketplace, elsewhere in the U.S. and for non-U.S. trade.

Figure 22: REMI Economic Forecasting Model – Basic Structure and Linkages



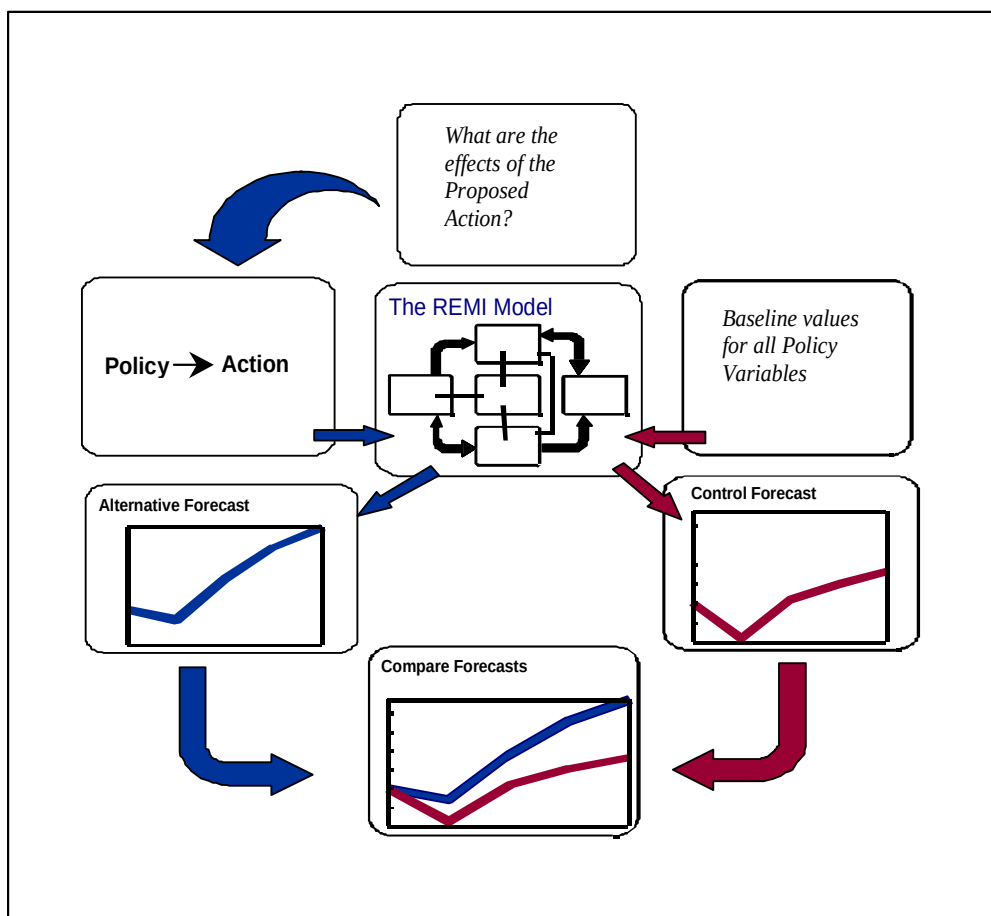
The REMI model identifies estimates of job impacts (and numerous other economic and demographic metrics) by comparing the base case³³ annual forecast using the above structure/feedbacks to the annual forecast when energy-related savings/costs or new dollars of investment are proposed through the *alternative* forecast. *Total* economic impacts result from the *direct* economic effects of SEP project investment. The *total* impact equals the *direct plus non-direct* impacts. Non-direct impacts are sometimes referred to as *ripple effect* in an economy. It is the presence of a comprehensive region-specific set of *multiplier effects* in the REMI economic simulation model that create additional economic responses once the *direct* effects have been introduced. Two economic mechanisms follow as a result of the direct program effects: *changes in Consumer demand* (often labeled '*induced*' effects) and *changes in*

³³ The regionally-calibrated software model is delivered with a *standard Regional Control* forecast out to 2050. This analysis has assumed that forecast is a sufficient long-term representation of the base case economies.

Intermediate demand or “B2B” (often labeled ‘indirect’ effects). The REMI model reports a *total* impact concept, and though it does not report separate *induced* and *indirect* contributions, both are accounted for, and we can segment these post-analysis.

The *total* economic impacts (stated in terms of jobs for this study objective) are expressed as a *difference* relative to jobs in year t without the program. Figure 23 portrays this relationship.

Figure 23. Identifying Economic Impacts in the REMI Framework



7.1.3 Translating SEP Project Direct Effects into Economic Events

The REMI model will translate the ways in which SEP dollars affect various segments through relevant *direct effects* that exert an influence on the local economy. Relevant direct effects include specific energy consumers (e.g. change in price, consumption or both), a region's economic self-sufficiency (by replacing imported purchases of energy generating feed stocks/energy driven components with more *locally* provided energy conserving devices/services), and the incremental cost required to achieve these goals. These *direct effects*, expressed as data inputs, will be developed as part of the team's data collection activities described in Section 4. In Figure 24 below, the left portion of the diagram portrays the set of direct effects that are possible with a broad range of energy-related investments/objectives. The major categories of direct effects associated with energy policies/investments and their potential to initiate macroeconomic responses are described below:

- Program operations (administrative) spending— dollars spent in operating the state's SEP program and paying incentives to business and household participants
- Household and business savings— dollar savings to businesses and households (resulting from reductions in energy and electric demand), realized as a result of the SEP funded project
- Household and business cost— additional household and business expenditures associated with the incremental cost of purchasing energy-efficient equipment/customer-sited RE systems (generally the total cost of new equipment minus incentives paid by the program and net of what would otherwise have been spent anyway). Could also include a ratepayer effect (*plus* as in lower rates/avoided costs or *minus* as in higher rates.)
- Other spending shifts— shifts in patterns of spending and business sales among sectors of the state's economy affecting the flow of dollars into, out of, and within the state. Included here are "*import substitution*" effects, new O&M spending requirements for new technology facilities/systems, as well as potential contraction for the power generating sector in light of energy-efficiency project uptake.

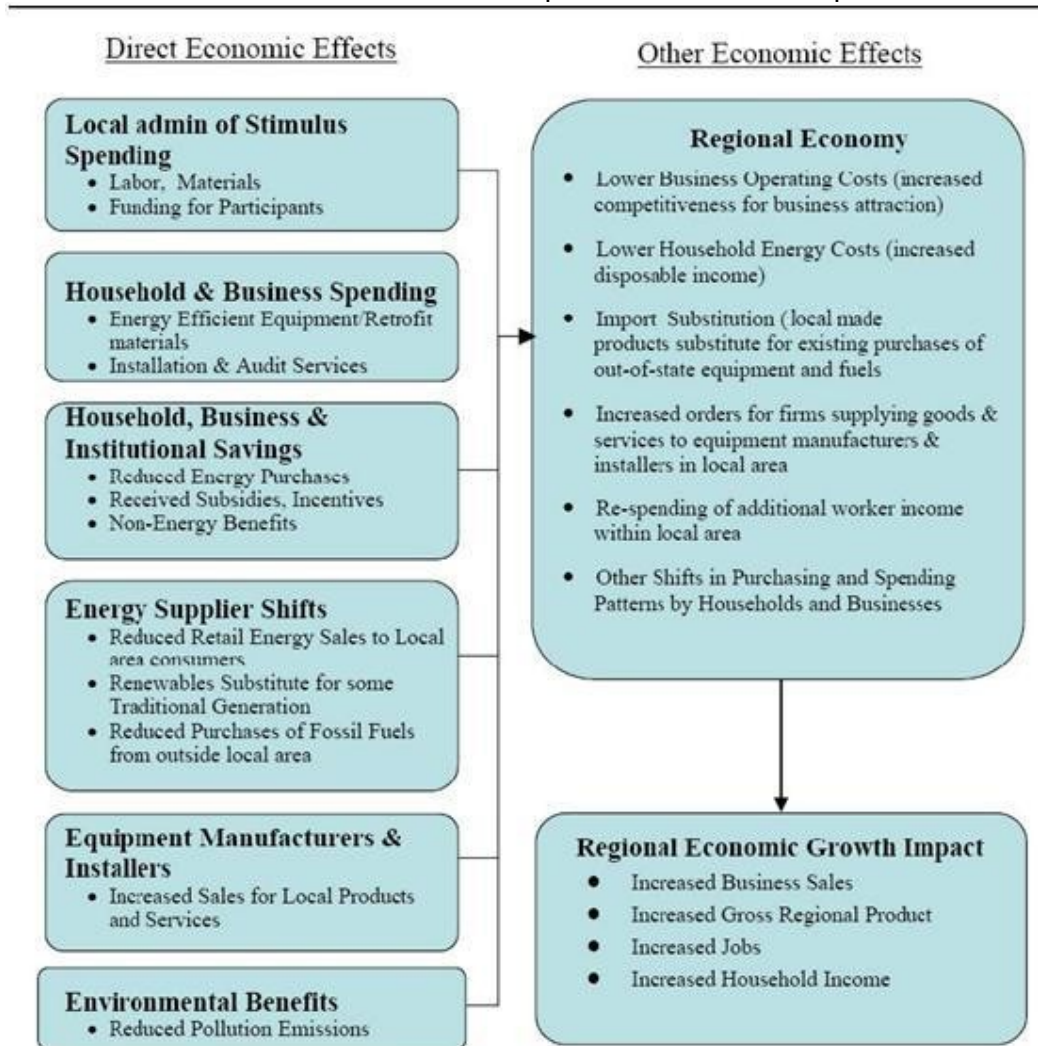
The "*mapping*" or translation of the above categories of direct effects into the economic impact model is depicted in the upper right portion of Figure 24. This entails careful delineation of instances when a new pattern of local demands arising from some or all energy customer

segments represents opportunities for greater reliance on “within region” sales, or none at all. The latter signals a continued import requirement albeit for an energy-efficient device instead of imported coal or petroleum feed stocks. Installation and other contractor services are more likely to be locally stimulated. Net savings to participating households and businesses (after paying off equipment investment cost differentials) have a clear pathway into the economic impact model and subsequent job creation.

Figure 24. REEM Framework for Energy Impact Analysis

[Renewable Energy Efficiency Mapping]

©2005-2011 Economic Development Research Group, Inc.



7.1.4 Presentation of Job Impacts

The key outputs of the macroeconomic modeling exercise will be presented to show the state-level job impact process at the BPAC or PA level. From the model's outputs, the KEMA team will be able to do the following:

- Distinguish the *time-phase* of impacts, e.g. short-term activities, long-term persistent changes,
- Distinguish the *direct* jobs from the *indirect* and *induced* job impacts,
- Use the results from attribution analyses by BPAC above to estimate the attributable job impacts associated with total project investment/implementation,
- Perform aggregations to harness BPAC/ state-level/national level job impacts from SEP projects by each program year to be evaluated.

8. Benefit-Cost Analysis

8.1 Types of Benefit-Cost Metrics

The primary benefit/cost analysis will be the SEP RAC test. This test calculates the lifetime source Btu saved per thousand dollars of SEP Recovery Act funding. This ratio is compared with the threshold of 10 million lifetime source Btu per \$1,000 of SEP funding. This threshold is the minimum level of energy savings each state's ARRA-funded SEP portfolio is required to target in its application. Moreover, guidance provided in the Funding Opportunity Announcement emphasizes that all ARRA-funded activities should ultimately meet this test over the life of the program.³⁴

Additional possible cost metrics suggested in the RFP for this study include:

- All energy cost savings reported in dollars and as a percentage of pre-treatment energy costs.
- Benefit/cost ratios reported as the net present value of cost savings as a proportion of total program expenditures.

However, the SEP guidance notes that, "the cost effectiveness test normally required within state regulatory environments that are focused on least cost net present value energy supplies do not apply to the SEP Recovery Act projects." Thus, tests commonly performed in the context of programs funded through utility rates will not be calculated. These tests include the Utility (or Program Administrator) Test, the Participant Test, the Total Resource Cost Test, and the Societal Test.

The Peer Review Panel has recommended that this study not conduct cost-effectiveness tests other than the RAC. "Given the limited funding available for evaluation and the needs identified earlier, and the complexity of developing cost-effectiveness definitions, collecting data and capturing all systems effects, the Panel believes that this would detract from the rigor needed to identify more critical outcomes resulting from SEP funding." At the same time, the Department of Energy has indicated a desire for a basic comparison of program benefits obtained with program costs,

³⁴ U. S. Department of Energy, National Energy Technology Laboratory, Financial Assistance Funding Opportunity Announcement, State Energy Program Formula Grants, April 24, 2009, Attachment 1, p. 28.

In light of this guidance, we plan to calculate only two indicators of cost-effectiveness:

1. The RAC test; and,
2. A basic benefit/cost ratio for each program period.

We will calculate these indicators at the national level for both the 2008 and ARRA portfolios of programmatic activities.

We will not calculate cost savings as a percent of pre-treatment energy costs. We do not expect to have good access to pre-treatment energy costs. Developing estimates of these quantities would require study resources that can be put to other uses.

We also will not attempt to produce a standard efficiency program cost-effectiveness test such as the Participant Test, Total Resource Cost Test, or Societal Test. One reason is the lack of applicability of such tests to SEP, as indicated above. Another reason is the challenge of obtaining comprehensive cost data, including customer expenditures, incremental equipment costs, as well as funding amounts from other sources. All these elements would be required to implement such tests.

Instead, our proposed benefit/cost ratio will compare the customer value (full costs paid by consumers) of energy savings attributable to SEP with the total SEP spending for each study period. This ratio is not directly comparable to conventional efficiency program benefit/cost tests. The proposed ratio compares the value of all realized benefits that are attributable to SEP with the SEP spending only, and does not take into account spending from leveraged funds or by the customer or recipient agency. The ratio will compare the dollar value of all savings induced by SEP with the SEP program costs. The published report will include a discussion of the rationale for presenting this ratio, and its lack of comparability to other common benefit/cost ratios.

Thus, the ratio calculated will be:

$$R = \sum_{sgpf} [E_{sgpf} P_{sgmfy} I_{sgpy} / (1+d)^{y-1}] / \sum_s D_s$$

where

R = overall national SEP benefit/cost ratio for the given study period

\sum_{sgpf} denotes summation over states s, segment g, BPAC group p, fuel f

E_{sgpf} = annual savings for state s, segment g, BPAC group p, fuel f

P_{sgpfy} = full consumer cost of energy for state s, segment g, BPAC group p, fuel f in year y

I_{sgpy} = dummy variable equal to 1 if BPAC group p, segment g has measure life greater than or equal to y, 0 otherwise

d = annual discount rate

D_s = total program spending on activities covered by this study, for state s during the study period

In this equation, the savings for each fuel and segment in each state will be calculated using savings factors determined from analysis of each BPAC-subcategory, applied to the spending by BPAC-subcategory.

8.2 Implementing Benefit-Cost Calculations

8.2.1 SEP RAC Test

For the SEP RAC Test, the information required is

1. Program spending.
2. Annual savings, in source Btu
3. Measure lifetime or Effective Useful Life (EUL).

Program spending is available from the program data already compiled. Annual energy savings will be determined as a primary product of the impact analysis. Translating the energy savings to source Btu requires multipliers for each fuel. For any factors that are not available from the site-specific data, we will rely on data from the Energy Information Administration (EIA).

Electricity is the energy source for which conversion to source energy is of most concern. One kWh delivered is equal to 3,412 Btu delivered. The source energy required to produce the kWh is approximately 10,000 Btu for fossil and nuclear power plants. The heat rate (Btu/kWh) depends on the plant type and efficiency. In most cases, the plant and fuels used to generate the electricity will not be known. We will use the average heat rate for U.S. plants, from EIA data.

For natural gas, 1 therm is equal to 100,000 Btu. 1 ccf is approximately equal to 1 therm. If natural gas consumption data are provided in ccf and the therm factors are not provided, we will apply the average therm factor from EIA data.

For bulk petroleum-based fuels including fuel oil, propane, and kerosene, consumption data are typically provided in gallons. We will use the EIA value of Btu per gallon.

Measure life estimates will rely on secondary sources. KEMA will conduct a literature review of available measure life estimates from various jurisdictions. Based on this review, we will establish a data base of measure life assumptions.

8.2.2 Net present value of energy savings versus program costs

To calculate this metric, in addition to the energy savings and program costs, we need to specify the energy prices and discount rate to be used in the net present value calculation.

For purposes of this calculation, we will calculate the dollar value of benefits over the life of the measures using EIA price data. EIA provides average retail price by sector and state for the current year, and also provides real consumer end use cost price projections for 25 years. We will use these data to establish retail prices for each state, sector, and year.

Following guidance from DOE, we will apply a discount rate based on OMB guidance per the annual update to "Circular No. A-94, "Guidelines and Discount Rates for Benefit-Cost Analysis of Federal Programs" which stands at 2.1 percent for a 20 year period in 2011. This is a standard basis for assigning discount rates for analysis of federal programs.

8.3 Level of Benefit-Cost Assessment

All of the above metrics can be calculated for each PA. However, there is no requirement that each PA be cost-effective by itself, by the SEP RAC test or other test.

The SEP-RAC test is intended to be applied to each state's portfolio. This study is not designed to evaluate directly any state as a whole. We propose to present the SEP RAC test and the net present value ratio at the national portfolio level, for each of the study periods.

9. Timeline

Figure 25 presents our current timeline for completing the SEP national evaluation. This project began with the initiation meeting held on October 18, 2010 at DOE's headquarters in Washington, DC.

Beginning in November 2010, the KEMA project team began receiving and reviewing versions of SEP program databases from DOE. First, we received third quarter 2010 data corresponding to the SEP ARRA period from DOE's PAGE database. We worked closely with DOE staff throughout November and December 2010 to ensure we had a complete dataset and that we understood the database contents and relationships.

Beginning in December 2010 and throughout January 2012, we received versions of PY2008 program tracking data from DOE's WinSaga database. This database was not as complete or as straight-forward as the data available for the ARRA period and, as a result, it took longer to complete our review.

During this same time frame, we began reaching out to DOE Project Officers to validate our understanding of the information contained in the various SEP databases. First, we focused on the ARRA period and, subsequently, we reached out to some of the more knowledgeable DOE Project Officers to gather information about the PY2008 period. In February 2011, we contacted NASEO Regional Coordinators and a few state program managers to fill in key gaps about the PY2008 period.

During February 2011, we also conducted several conference calls with the ORNL, DOE and KEMA project team to discuss our preliminary approach for evaluating some of the key outcomes of the program, including:

- Gross energy savings definitions and calculational approaches
- Environmental and employment impact evaluation methods
- Attribution framework and evaluation approaches

In addition, the KEMA team met its goal to submit the 60 day ICR notice by March 28, 2011.

The following summarizes the timeline for completing some of the remaining high-level evaluation activities:

- **Obtain OMB approval for evaluation data collection materials.** As mentioned earlier, there is a strict, lengthy approval process for evaluation surveys and other data collection materials. We anticipate approval sometime in the fall of 2011. We have estimated November 1, 2011 but will work closely with DOE to speed up the process as much as possible. Currently, the KEMA team is pursuing an Emergency ICR submission process which may expedite study implementation.
- **Sample selection.** We anticipate having our final sample selected in late June 2011.
- **Implement evaluations for sampled PAs.** During the summer of 2011, we will be working on developing common evaluation protocols (i.e., attribution approaches for specific types of PAs), as well as various evaluation “tools” (i.e., energy savings calculators, CO2 models, labor market models, etc.). Survey instrument development has begun in May 2011, with pretests scheduled for July 2011, and finalized materials ready for OMB submittal by August 2011. Large sample data collection cannot begin until the data collection materials are approved by OMB and, thus, we scheduled this to begin in early November 2011 and go through March 2012. Evaluation analyses of the all studies will begin in February 2012 through July 2012.
- **Project reporting.** In addition to weekly meetings and monthly reports, we have scheduled a number of interim report deliverables to provide ORNL and DOE with more timely feedback on the progress and early results from our overall evaluation effort. We have scheduled the following interim reports:
 - o July 1, 2011 (coincident with the final sample design milestone, drafting of data collection instruments, and ongoing development of evaluation protocols)
 - o November 7, 2011 (coincident with expected final OMB approval and completed analyses of medium-low rigor evaluation analyses)
 - o April 2, 2012 (coincident with expected completion of large sample data collection)

In addition, the timeline allows for a draft final evaluation report to be completed by end of July 2012 with the revised final evaluation report completed September 17, 2012.

Figure 25. SEP National Evaluation Timeline

	Q3 2010												Q1 2011												
	18-Oct-10	25-Oct-10	1-Nov-10	8-Nov-10	15-Nov-10	22-Nov-10	29-Nov-10	6-Dec-10	13-Dec-10	20-Dec-10	27-Dec-10	3-Jan-11	10-Jan-11	17-Jan-11	24-Jan-11	31-Jan-11	7-Feb-11	14-Feb-11	21-Feb-11	28-Feb-11	7-Mar-11	14-Mar-11	21-Mar-11	28-Mar-11	
Project Initiation Meeting (October 18, 2011)	m																								
Evaluation Planning																									
Received program data from DOE																									
Contacted DOE/NASEO/SEP staff																									
Conference calls to discuss evaluation methods, OMB approval process, etc.																									
Draft evaluation plan submitted to ORNL, DOE and Peer Review Panel																									
Peer Review Panel Meeting																					m				
Incorporate ORNL, DOE and Peer Review Panel Comments																									
Final evaluation plan																									m

	Q1 2011				Q2 2011								Q3 2011					Q4 2011																						
	28-Feb-11	7-Mar-11	14-Mar-11	21-Mar-11	28-Mar-11	4-Apr-11	11-Apr-11	18-Apr-11	25-Apr-11	2-May-11	9-May-11	16-May-11	23-May-11	30-May-11	6-Jun-11	13-Jun-11	20-Jun-11	27-Jun-11	4-Jul-11	11-Jul-11	18-Jul-11	25-Jul-11	1-Aug-11	8-Aug-11	15-Aug-11	22-Aug-11	29-Aug-11	5-Sep-11	12-Sep-11	19-Sep-11	26-Sep-11	3-Oct-11	10-Oct-11	17-Oct-11	24-Oct-11	31-Oct-11	7-Nov-11			
Project Initiation Meeting (October 18, 2011)																																								
OMB Approval of Data Collection Materials					m																																			
Preparation and Required Comment Periods																																								
OMB Approval Period																																								
OMB Approval (estimate)																																								m

